

Proceedings of the Northeast Decision Sciences Institute



March 28-30, 2008



Program Chair
Snehamay Bannerjee
Rutgers University - Camden

Proceedings Editor
John Affisco
Hofstra University

PERSONALITY AND ACCOUNTING: DOES A FINANCIAL APTITUDE EXIST?

William E. Bealing, Jr., Bloomsburg University, (570) 389-4386, wbealing@bloomu.edu
A. Blair Staley, Bloomsburg University, (570) 389-4392, astaley@bloomu.edu
Richard L. Baker, Bloomsburg University, (570) 389-4561, rbaker@bloomu.edu

ABSTRACT

This study examines the causal relationship between personality type and performance in an entry level accounting course. The effect of specific personality scores from the Keirseley Temperament Sorter personality preferences test is regressed on grades obtained by non-accounting business majors who enrolled in the introductory accounting courses at one mid-sized, public university. The results indicate an individual's personality preference in one, specific, categorical type—namely, the *sensing* personality type—is positively related to their grade in an introductory accounting course.

INTRODUCTION

Historically, particularly strong aptitudes demonstrated by certain individuals have been ascribed to their unique personality types. For instance, most individuals would accept the premise that certain members of society possess a unique aptitude for music.

Previous work by Jackson and Lawty-Jones [7] and Honey and Mumford [6] have linked learning styles to a student's personality, although Loo [10] pointed out that learning styles and personality are not actually the same constructs. While a significant body of research exists concerning the relationship of student learning styles to academic performance, far less has been written about the relationship between student personality type and classroom performance.

This study investigates the premise that a measurable correlation exists between identifiable personality types and academic performance in a specific academic discipline: accounting.

BACKGROUND

The theory that individuals exhibit specific personality types was introduced by Carl Gustav Jung in his book *Psychological Types* in 1921. In his book, Jung documented a set of typical differences in the human psyche that he termed "introverted" and "extroverted." He posited that extroverted individuals prefer the outer world of people and things while introverted individuals prefer the inner world of ideas and concepts. Jung also posited that psychological types were related to certain basic psychological functions which he termed thinking, feeling, sensation, and intuition [17].

Jung's descriptions of introversion and extroversion represent the extremes of human behavior. Each individual will actually display a tendency to lean more in the direction

of introversion or extroversion but possess both qualities to some degree. In actuality, each individual possesses all of the attributes of sensation, thinking, feeling, and intuition, but demonstrates a predisposition to rely on some attributes more than others. The individual's predisposition toward individual attributes determines their specific personality type.

In 1936, Jung published a related paper that he described as an appendix to his original book, *Psychological Types*. In that paper, he described in greater detail his four basic psychological types; thinking, feeling, sensation, and intuition. He proposed that sensation establishes that something exists, thinking establishes what it means, feeling attaches a value to it, and intuition determines when it comes and whether it will last [17].

It has been postulated that individuals perceive their environment through either a sensing or an intuitive paradigm [16]. Sensing types tend to organize input sequentially, in a step-by-step process. Conversely, intuitive types start with a top down view of broad concepts. They then take these abstract concepts and organize them as a more workable general framework. Intuitive types also tend to dislike activities that are detail oriented.

On the other hand, an individual's decision making processes are dominated by either thinking or feeling. Those individuals who conform to the thinking type definition tend to use a logical, objective decision process, while those who resemble the feeling type are inclined to use a value-based or subjective process which puts more emphasis on how the decision will impact others [16].

Finally, an individual's attitude toward their external environment and the structure of their everyday living is dominated by judging or perceiving. Those individuals who conform to the judging-type definition prefer to plan ahead and prefer closure and the settling of things. Those individuals who conform to the perceiving-type definition prefer to postpone decisions and keep things open-ended.

Jung also drew some notable caveats regarding psychological typing. First, the basic typologies of introversion, extroversion, sensation, thinking, feeling, and intuition were never intended to encompass all aspects of human personality. Furthermore, the purpose of psychological typing is not to classify people into neat, restrictive categories. He maintained that simply categorizing individuals, in itself, would be a meaningless exercise. Jung believed that psychological typing merely provided the methodology for conducting empirical research by providing points of view and guidelines that reduce the plethora of individual personality types down to a manageable array. In addition, psychological typing may aid us in understanding the broad differences in personality and personal preferences that occur among individuals [17].

HYPOTHESIS

Hypotheses: There is a significant relationship between personality type and performance in an introductory accounting course.

THE RESEARCH INSTRUMENT

Research into the area of learning styles has been ongoing. This has resulted in the development of over twenty different currently available models [4] [15] [5].

Much of the previous research into the link between personality type and accounting has been based on the Myers-Briggs Type Indicator (MBTI). This is a formal preferences based survey instrument consisting of over 120 questions designed to access an individual's personality preferences in four primary areas: introvert/extrovert, sensing/intuitive, thinking/feeling, and judging/perceiving. In theory, every individual has a predisposed preference to one of each of the above pairs. Thus, there is a potential of sixteen (four squared) individual personality types and, theoretically, each individual can be categorized as one of those types.

Wheeler [18] found 16 published accounting research articles using the MBTI instrument. Of the 16 articles, four examined the relationship between academic performance of accounting students and their indicated MBTI personality type. Specifically, Ott, et al. [14] found that individuals categorized as sensing (S) and thinking (T) performed better, as measured by course grades, in courses using the lecture method.

Utilizing the MBTI instrument, Nourayi and Cherry [11] examined the relationship between the performance of several students in seven accounting courses and their individual personality preferences. The only significant relationship found was that students categorized as sensing (S) outperformed intuitive (I) students in three (Tax, Auditing and Intermediate II) of the seven courses analyzed in their study.

While not examining the linkage between personality type and course grade, Landry et al. [9] focused on computer usage and the MBTI personality of accounting students. They found that the STJ-preference (sensing, thinking, judging) personality type was over represented, comprising 42% of their entire sample, with ISTJ (introverted STJ's) comprising 17% and ESTJ (extroverted STJ's) comprising 25% respectively.

Oswick and Barber [13] examined the MBTI personality preferences of undergraduate non-accounting majors and contrasted the individual findings to each student's performance in an introductory accounting course, as measured by course grades. They found that there were no statistically significant correlations between indicated personality preference and performance.

Keirseley and Bates [8] developed their own independent personality preferences test utilizing the same 16 personality types as the MBTI instrument. Since its inception, the Keirseley Temperament Sorter (KTS) has grown to become the most widely used personality inventory measuring tool in the world, surpassing even the MBTI instrument.

The authors have chosen to use the KTS instrument in their research for several reasons. First, the 16 resulting personality classifications used by the KTS instrument correspond directly to those of the MBTI instrument that had been used in the previous research

conducted in the field. Second, it can be administered and scored by the researchers in approximately twenty minutes without having to send the instrument off site. In addition, the authors have complete access to the individuals' responses to each question, not just their overall personality type score. This has proved to be invaluable in facilitating a far more detailed analysis of the results. Finally, the cost of administering the MBTI instrument to several hundred students proved to be prohibitive while the cost of administering the KTS instrument fell within the researchers' budget constraints. Specifically, the authors propose that the four primary personality types—introvert/extrovert, sensing/intuitive, thinking/feeling, and judging/perceiving—are related to success in an introductory accounting course.

METHODOLOGY

Over a three-year period, the authors administered the KTS instrument to 95 non-accounting business majors enrolled in the basic financial accounting course for non-accounting majors. To motivate the students to conscientiously complete the survey, it was administered during a normal class period. In addition, the students were informed that they could receive access to their individual personality profiles, as well as a copy of a common careers list for each personality type, simply by contacting the faculty member who administered the survey instrument. They were also told to leave the questionnaire blank if they chose not to participate. No blank forms were received. The data were coded centered on zero so that the four primary areas of interest – introvert/extrovert, sensing/intuitive, thinking/feeling, and judging/perceiving – could be measured with four variables.

In compliance with university policy, the research was pre-approved by the University's Institutional Review Board and the students were informed that, although their individual results would remain strictly confidential, aggregated research results would be included in various future research findings made public by the authors.

RESULTS

Means, standard deviations, correlations, and reliabilities

Table 1 shows the means, standard deviations, correlations, and internal reliabilities for the variables of the study. Only the sensing/intuitive variable was statistically significant with respect to the variable success in an accounting course ($p < .05$). The sensing/intuitive variable was also statistically significant with respect to the thinking/feeling and judging/perceiving variables. The thinking/feeling variable was significant with respect to the judging/perceiving variable. The measures had satisfactory internal reliabilities, with the exception of the thinking/feeling measure, which had an alpha coefficient (.66) that is marginally below the .70 level recommended by Nunnally [12].

Table 1
Means, Standard Deviations, Correlations, and Reliabilities^a

Variable	Mean	SD	1	2	3	4	5
1. Success in Accounting Course	1.91	1.07	(N/A)				
2. Thinking (-) or Feeling (+)	0.37	3.54	-.08	(.66)			
3. Sensing (-) or Intuitive (+)	-2.97	3.43	-.22*	.25*	(.70)		
4. Extrovert (-) or Introvert (+)	-1.23	2.54	.11	-.07	-.12	(.73)	
5. Judging (-) or Perceiving (+)	-2.37	3.75	-.14	.33**	.58**	-.10	(.74)

*, ** Significant at the .05 and .01 levels, respectively.

^a N = 95. Alpha internal reliability coefficients are in parentheses on the diagonal.

Primary analysis

The model was tested with conventional multiple regression analysis using the AMOS [1] computer program. The analysis focused on the proposed model specifying relationships between the four latent variables of interest and the observed variable of interest. For purposes of the analysis, introvert/extrovert, sensing/intuitive, thinking/feeling, and judging/perceiving were treated as correlated exogenous variables, which statistically controls for the interrelationship between the four variables when the structural path coefficients in the model are estimated [2].

The standardized coefficients for the hypothesized direct paths between the variables in the proposed model, and the correlation between the exogenous variables were obtained. Of primary concern, the only coefficient for the hypothesized direct paths between the variables in the proposed model that was significant (at $p < .1$) was sensing. Hypothesis 1, which proposes a relationship between personality type and success in an accounting course, was supported by the significant direct path between sensing and grades in the course.

IMPLICATIONS

The results generated from this research project indicate that there a relationship between a sensing (S) personality preference, as measured by the KTS instrument, and performance in an introductory accounting course. The results appear to support the premise that a student with a strong (S) component had the best chance for success in the course. This is not surprising given that a basic accounting course tends to be taught in a logical, step by step, sequence and involves the processing of detailed information. These are exactly the types of activities for which sensing (S) individuals normally have a preference. Students who are naturally intuitive could be counseled that accounting may not come easily for them and on what steps may be necessary for these students to do to succeed in the course. Such counseling may not only help the student succeed, but also may help the instructor succeed, as the instructor would be aware of the students who may need extra attention and assistance.

SUMMARY AND FUTURE RESEARCH

The idea that certain individuals are predisposed to having an "accounting ability" was examined. The desired personality type for an undergraduate accounting student appears to be an individual who possesses a strong (S) personality preference. To lend credence to this finding, the research needs to be replicated at several different universities of diverse sizes as well as within business colleges with varying reputations regarding the quality and rigor of their programs.

In addition, the participants in this study need to be tracked throughout their academic career to determine whether the relationships remain consistent as the students take progressively more advanced accounting courses. Simultaneously, research into the issue of whether an "S" individual has a tendency to do better in other quantitative courses, such as economics or statistics, needs to be conducted. The question of whether "S" types do better in courses in other business disciplines should also be examined. If these findings are able to be generalized, there are obvious potential implications for the admissions criteria for accounting majors and possibly all business majors.

Also, if a statistically significant number of observations could be amassed, relationships between all 16 possible personality types and performance in accounting courses should be examined. Finally, if a statistically significant number of observations could be amassed, relationships between all 70 KTS questions and performance in accounting courses should be examined.

References available upon request from William Bealing.

A PRELIMINARY ANALYSIS OF METHODS USED BY ACCOUNTING PROFESSORS TO PREVENT DISHONEST BEHAVIOR

David E. Morris, 134 Newton Oakes Center, North Georgia College & State University,
Dahlonega, GA 30597, (706) 864-1611, dmorris@ngcsu.edu

ABSTRACT

Recent research has indicated that a large percentage of college students have admitted to engaging in academically dishonest (cheating) behavior at some point during their college careers. This has resulted in a great deal of research related to this topic. Many techniques have been suggested as ways to decrease this troubling trend. The primary purpose of this paper is to describe a study which surveyed accounting professors asking if they were employing some of the more commonly prescribed techniques and if they believed the use of these methods had a positive effect on the problem with cheating. A survey was sent to 600 accounting professors. The results indicated that a significant number of the respondents were using at least some of these methods. However, the majority indicated that they believed the use of such methods had, at best, only limited benefits.

INTRODUCTION

In recent years there have been numerous research studies confirming that cheating is a very serious problem in colleges and universities in this country (and around the world). This is especially of concern to accounting educators whose students' careers are dependent upon public trust and confidence. Morris and Kilian [4] have indicated that accounting majors admitted to having cheated just as often as other business majors. This concern of all college educators has resulted in a lot of research and discussion as to how to best decrease the amount of cheating that takes place in college classrooms across the country.

A review of the literature by the author identified several techniques that are frequently mentioned as being efficient in reducing the amount of cheating that takes place on our campuses. For example, Davis [1] suggests that, in assigning papers, professors make very clear the differences between plagiarism, paraphrasing and direct citation. Rocklin [5] suggests that professors require interim evidence of students' work in writing papers. Dodd [2] states the importance of a school having a written academic policy and that the policy be included in the class syllabus. It has also been suggested by Kerkvliet and Sigmund [3] that cheating might be decreased by professors using different versions of the same exam.

This study sought to find out if accounting educators are using some of the more popular methods recommended and if they believe the use of these methods has had the desired effect of decreasing the amount of cheating in their classes.

Please note that in this paper cheating and dishonest behavior are considered synonymous.

HYPOTHESIS DEVELOPMENT

A significant amount of literature has been written on methods to decrease cheating on college campuses. Given the concern accounting educators have with this problem, it was believed by the author that a significant number of respondents would indicate that they did use some of these methods as a way to help prevent cheating in their classes. This led to the first hypothesis:

H1: A significant number of accounting professors will indicate that they do employ some of the methods stated in the survey as ways to help decrease the amount of cheating in their classes.

A review of the literature reveals very little evidence as to whether these methods used by accounting professors are effective and to what degree they might be effective. Based on the author's professional experience and in discussions with colleagues it was believed (by the author) that a significant number of accounting professors would indicate they have had only limited success in decreasing the amount of cheating in their classes. This led to the second hypothesis:

H2: A significant number of accounting professors will indicate only limited success in the use of these methods to decrease cheating in their classes.

SAMPLE SELECTION AND INSTRUMENT

In this initial study a sample of 600 accounting professors were selected at random from *Prentice Hall 2006-2007 Accounting Faculty Directory* and sent (via email) a survey asking them, among other things, if they used various methods cited in the literature as ways to reduce the amount of cheating in their classrooms. They were then asked if they believed the use of these methods had been effective in reducing dishonest behavior.

RESULTS

The results can be seen in Table 1. The first hypothesis stated a significant number of accounting professors would indicate that they do employ some of the methods that have been recommended to reduce cheating in their classes. The results indicate that the majority of the respondents did employ some of the methods. Thus, the first hypothesis was supported.

The second hypothesis stated that a significant number of respondents would indicate only limited success in the use of these methods to reduce cheating in their classes. The results support the hypothesis. Of the respondents who used at least one of these methods, only 37.9% indicated they believed it had a significant effect in reducing cheating.

TABLE 1
Responses of Accounting Professors
(stated in percentages; n=56)

Responses regarding uses of various methods to reduce cheating in the classroom

1.	Does your school have a written Academic Integrity Policy?	Yes:	96.7%
		No:	3.3%
		I do not know:	0.0%
2.	Do you include the school's Academic Integrity Policy in your syllabus?	Yes:	83.3%
		No:	13.3%
	My school does not have a written policy:		3.3%
3.	Do you explain to your students what you consider to be acceptable and unacceptable behavior with regards to academic honesty?	Yes:	86.2%
		No:	13.8%
4.	If you give multiple-choice questions on your exams, do you give different versions of the same exams by rearranging the questions and/or the answers?	Yes:	63.3%
		No:	20.0%
	I do not give multiple choice questions:		16.7%
5.	When you give exams, do you require students to alternate seats?	Yes:	43.3%
		No:	56.7%
6.	Do you administer exams on the "honor system"?	Yes (always):	3.4%
		No (never):	65.5%
		Sometimes:	31.0%
7.	Do you tell students that papers they submit may be submitted to software programs that are designed to detect plagiarism?	Yes:	46.7%
		No:	33.3%
	I do not require students to write papers:		20.0%
8.	Do you require students, when writing papers, to submit interim evidence of their process?	Yes:	20.7%
		No:	58.6%
	I do not require students to write papers:		20.7%

9. When assigning papers, do you try to clarify the differences between plagiarism, paraphrasing, and direct citation?
- | | |
|--|-------|
| Yes: | 63.3% |
| No: | 16.7% |
| I do not require students to write papers: | 20.0% |
10. Do you state your policy regarding the submission of a paper in your class that has been previously submitted in another class?
- | | |
|--|-------|
| Yes: | 34.5% |
| No: | 44.8% |
| I do not require students to write papers: | 20.7% |
11. If you do employ one or more of these methods, which statement best reflects the effectiveness of those methods in deterring cheating in your classes?
- | | |
|---|-------|
| I believe the use of these methods has a significant effect: | 37.9% |
| I believe the use of these methods has had only a limited effect: | 41.4% |
| I believe the use of these methods has had not effect: | 0.3% |
| I do not use any of these methods: | 10.3% |

Additional information obtained from the survey

12. With regards to dishonest behavior in your classes:
- | | |
|--|-------|
| Instances of such behavior have increased in recent years: | 40.0% |
| Instances of such behavior have decreased in recent years: | 6.7% |
| Instances of such behavior have remained about the same: | 53.3% |
13. Do you believe the use of the Internet in YOUR classes has:
- | | |
|---|-------|
| Contributed to increases in the instances of dishonest behavior: | 50.0% |
| Contributed to decreases in the instances of dishonest behavior: | 0.0% |
| Had no significant effect on the instances of dishonest behavior: | 50.0% |
14. Do you believe your school will support you if you accuse a student of cheating?
- | | |
|--|-------|
| Yes: | 70.0% |
| No: | 13.3% |
| I do not know if my school would support me: | 16.7% |

REFERENCES

- [1] Davis, Barbara . “Preventing Academic Dishonesty” in *Tools for Teaching*, Jossey-Bass, 1993, [accessed at <http://teaching.berkeley.edu/bgd/prevent.html>].
- [2] Dodd, T. “Honor Code 101 An Introduction to the Elements of Traditional Honor Codes, Modified Honor Codes and the Academic Integrity Policies,” 2007, [accessed at www.academicintegrity.org].

- [3] Kerkvliet, J and Sigmund, C. "Can We Control Cheating in the Classroom?" *Journal of Economic Education*, Fall 1999, 30(4), 331-343.
- [4] Morris, D. and Kilian, C. "Do Accounting Students Cheat? A Study Examining Undergraduate Accounting Students' Honesty and Perceptions of Dishonest Behavior," *Journal of Accounting, Ethics & Public Policy*, 2006, 5(3), 375-386.
- [5] Rocklin, Tom. "Downloadable Term Papers: What's a Prof. to Do?" Teaching Center at the University of Iowa, 1996, [accessed at www.uiowa.edu].

EFFECTS OF RECENT LEGISLATION REGARDING HUSBAND-WIFE PARTNERSHIPS FOR TAX PURPOSES

Nathan Oestreich, 619-594-2478, drno@sdsu.edu
Carol Venable, 619-594-2662, Carol.Venable@sdsu.edu
James E. Williamson, 619-594-6021, James.Williamson@sdsu.edu
School of Accountancy, College of Business Administration, San Diego State
University, 5500 Campanile Drive, San Diego, CA 92182-8221,

ABSTRACT

The *Small Business and Work Opportunity Tax Act of 2007* provides that a married couple who operate an unincorporated enterprise, that would otherwise be a partnership for tax purposes, as joint venturers can elect to treat the entity as two sole proprietorships, rather than a partnership. Couples can review and update the way they currently share profits and losses, and even change the way they share responsibilities so as to change the allocations. This allows couples to exercise significant control over the amount of self-employment tax they pay, and under current Social Security rules, the amount of Social Security benefits they receive. This study reviews the history of husband-wife partnerships and sole proprietorships and how the spouses can affect the amount of their anticipated Social Security benefits.

INTRODUCTION

Married couples can hold title to businesses that they own in a variety of ways, including corporations, partnerships, and sole proprietorships. In the case of corporations, the owners can also be employees. If they operate as a sole proprietorship, the non-owner spouse can be employed, and the owner spouse is self-employed. If the spouses are partners, they are each self-employed and neither can be treated as an employee.

The *Small Business and Work Opportunity Tax Act of 2007* [3] allows couples who would be considered partners to elect to report their business activities as two separate sole proprietorships. This allows husbands and wives to manipulate the amount of self-employment tax they pay, and as a result, their benefits. The same results can be accomplished using a partnership, but the new law allows them to avoid the filing requirements and the complexities of the partnership form.

WHO IS A PARTNER?

The definition of a partner is found in Internal Revenue Code [IRC] § 7701(a) (2), which reads as follows:

Partnership and partner. The term “partnership” includes a syndicate, group, pool, joint venture, or other unincorporated organization, through or by means of which any business, financial operation, or venture is carried on, and which is not, within the meaning of this title, a trust or estate or a corporation; and the term “partner” includes a member in such a syndicate, group, pool, joint venture, or organization.

The technical definition partners and partnerships had evolved over many years, but effectively changed when the check-the-box regulations were issued on December 18, 1996 [8], which were effective January 1, 1997. These regulations usually allow unincorporated business entities to choose whether they will be treated as a corporation or partnership for tax purposes. To implement the regulations, a new set of rules needed to be developed. First, Regulation 301.7701(a) stipulates that the question as to whether an organization is separate from its owners for federal tax purposes is determined independently of its status under local law. Then, it goes on to provide that a “. . . joint venture or other contractual arrangement may create a separate entity for federal tax purposes if the participants carry on a trade, business, financial operation, or venture and divide the profits therefrom.”

In order to implement the check-the-box procedure, new concepts were adopted. First, the entity must be a business entity [5], and secondly, it must be an eligible entity [6]. Under the regulations a “business entity” is a federal tax entity that is not classified as a trust or otherwise subject to special treatment. Essentially, a “business entity” is one that has an owner or owners carrying on a trade or business for profit. An “eligible entity” is any business entity not categorized as a “corporation.” For example, a corporation formed under state law is not an eligible entity.

Eligible entities (e.g., LLCs, LLP’s, partnerships) may simply select their form for tax purposes. Newly formed eligible entities fall into a default category. Those with one owner are disregarded; more than one owner, partnerships. However, they may check-the-box and elect to be treated as corporations.

HUSBAND-WIFE BUSINESSES

Properly classifying an entity that is owned and operated by a wife and/or her husband can be difficult. Is it a partnership, a sole proprietorship, a joint venture or none of the above? As a result, couples have typically had significant discretion, depending on how they organized and operated their business.

One possibility is that the business is owned and operated by one spouse, who may or may not compensate the other spouse as an employee, or when appropriate, as a consultant. In the case of wages, they are deductible and appropriate payroll withholding and reporting are required. If the couple files jointly, there is little income tax effect, but there are significant payroll tax and self-employment tax differences. There are the fringe benefit possibilities for the employee.

If the couple elects to form a partnership (or simply report as partners), they have significant discretion as to how they hold title and how they allocate profits and losses. Neither is an employee for tax purposes, and neither can benefit from employee fringe benefit programs.

Couples who elect corporate status (either by incorporating the business or by checking-the-box) also have significant discretion. Either the wife or the husband, or both, can be employed by the business and qualify for reasonable salaries and benefits. Appropriate withholding must occur and the requisite payroll reporting made. If the S election is made, only employees with 2 percent or less ownership qualify for employee fringe benefits.

Each form of business provides possible employee fringe benefits. Traditionally, the medical benefit was a key element in entity selection. However, the owner of an unincorporated business (or S corporation) who provides medical coverage for her/his employees can deduct the cost of her/his medical plan as a deduction for adjusted gross income.

WIFE-HUSBAND PARTNERSHIPS

Couples have some control over whether their businesses are partnerships for tax purposes, but caution is necessary so as to avoid unintended consequences. One way to avoid partnership status is to avoid joint ownership. Perhaps the clearest statement of IRS policy on wife-husband ownership appears in Publication 541 (under frequently asked questions) [2].

Can a husband and wife run a business as a sole proprietor or do they need to be a partnership? It is possible for either the husband or the wife to be the owner of the sole proprietor business. When only one spouse is the owner, the other spouse can work in the business as an employee. If the spouses intend to carry on the business together and share in the profits and losses, then they have formed a partnership.

Similar, but stronger, wording is found in IRS Publication 541: Partnerships [2].

Husband-wife partnership. If spouses carry on a business together and share in the profits and losses, they may be partners whether or not they have a formal partnership agreement. If so, they should report income or loss from the business on Form 1065.

Community ownership, however, is problematic under these guidelines. In order to provide couples some certainty as to their status, the government issued Rev. Proc. 2002-69. This guidance provides that the IRS will accept the treatment of a business that is owned solely by a husband and wife as a disregarded entity or a partnership as indicated by the owners' actions. They can file as sole proprietor(s) or as partners.

LEGISLATIVE CHANGE

Congress must have believed that further clarification and flexibility was appropriate, as in the *Small Business and Work Opportunity Tax Act of 2007* [3] amended § 761 (definitions), amending and redesignating subsection (f) as subsection (g) and inserting the following:

(f) Qualified Joint Venture-

- (1) IN GENERAL- In the case of a qualified joint venture conducted by a husband and wife who file a joint return for the taxable year, for purposes of this title--
 - `(A) such joint venture shall not be treated as a partnership,
 - `(B) all items of income, gain, loss, deduction, and credit shall be divided between the spouses in accordance with their respective interests in the venture, and
 - `(C) each spouse shall take into account such spouse's respective share of such items as if they were attributable to a trade or business conducted by such spouse as a sole proprietor.

- (2) QUALIFIED JOINT VENTURE- For purposes of paragraph (1), the term 'qualified joint venture' means any joint venture involving the conduct of a trade or business if--
- (A) the only members of such joint venture are a husband and wife,
 - (B) both spouses materially participate (within the meaning of section 469(h) without regard to paragraph (5) thereof) in such trade or business, and
 - (C) both spouses elect the application of this subsection.'

The statute also addresses directly how this election applies to the determination of the self-employment tax as follows:

(b) Net Earnings From Self-Employment-

- (1) Subsection (a) of section 1402 (defining net earnings from self-employment) is amended by striking ', and' at the end of paragraph (15) and inserting a semicolon, by striking the period at the end of paragraph (16) and inserting `; and', and by inserting after paragraph (16) the following new paragraph:
(17) notwithstanding the preceding provisions of this subsection, each spouse's share of income or loss from a qualified joint venture shall be taken into account as provided in section 761(f) in determining net earnings from self-employment of such spouse.'
- (2) Subsection (a) of section 211 of the Social Security Act (defining net earnings from self-employment) is amended by striking 'and' at the end of paragraph (14), by striking the period at the end of paragraph (15) and inserting `; and', and by inserting after paragraph (15) the following new paragraph:
(16) Notwithstanding the preceding provisions of this subsection, each spouse's share of income or loss from a qualified joint venture shall be taken into account as provided in section 761(f) of the Internal Revenue Code of 1986 in determining net earnings from self-employment of such spouse.'

These changes apply to taxable years beginning after December 31, 2006.

In their explanation of this provision, the Staff to the Joint Committee on Taxation, after explaining the law, close with the following statement: "The provision is not intended to prevent allocations or reallocations, to the extent permitted under present law, by courts or by the Social Security Administration of net earnings from self-employment for purposes of determining Social Security benefits of an individual" [4].

EFFECT OF EARNINGS ON SOCIAL SECURITY BENEFIT

There is currently a direct, although not necessarily proportional, relationship between one's reported earnings the eventual social security benefit she/he receives. The benefit is based one's best 35 years of reported earnings, adjusted for inflation. The retirement benefit, called the primary insurance amount, is described in the Social Security Administration website (SSS.gov) as follows [7] (This research does not consider the effect on the taxation of the Social Security Benefit under § 86):

The "primary insurance amount" (PIA) is the benefit (before rounding down to next lower whole dollar) a person would receive if he/she elects to begin receiving retirement benefits at his/her normal retirement age. At this age, the benefit is neither reduced for early retirement nor increased for delayed retirement. PIA formula bend points The PIA is the sum of three separate percentages of portions of average indexed monthly earnings. The portions depend on the *year* in which a worker attains age 62, becomes disabled before age 62, or dies before attaining age 62.

For 2007 these portions are the first \$680, the amount between \$680 and \$4,100, and the amount over \$4,100. These dollar amounts are the "bend points" of the 2007 PIA formula. See table showing bend points for years beginning with 1979 (table also includes bend points in maximum family benefit formula).

PIA Formula

For an individual who first becomes eligible for old-age insurance benefits or disability insurance benefits in 2007, or who dies in 2007 before becoming eligible for benefits, his/her PIA will be the sum of:

- (a) 90 percent of the first \$680 of average indexed monthly earnings, plus
 - (b) 32 percent of average indexed monthly earnings from \$680 to \$4,100, plus
 - (c) 15 percent of his/her average indexed monthly earnings over \$4,100.
- Rounded to the next lower multiple of \$.10.

Determination of the PIA Bend Points for 2007

Amounts in formula	<u>Average wage indices</u>		Bend points for 1979	
	For 1977:	9,779.44	First:	\$180
	For 2005:	36,952.94	Second:	\$1,085
Computation of bend points for 2007	<u>First bend point</u> \$180 times \$36,952.94 divided by \$9,779.44 equals \$680.15, which rounds to \$680.		<u>Second bend point</u> \$1,085 times \$36,952.94 divided by \$9,779.44 equals \$4,099.82, which rounds to \$4,100.	

Planning

Couples jointly owning businesses have varying goals with regard to the Social Security system. Some, usually younger individuals, prefer to pay the minimum possible [1]. Others, usually older individuals, may prefer to pay more tax to enhance their eventual benefits. But, planning is not so simple as considering age, as the amount of reported wages can effect disability and survivor benefits. Probabilities involved in marital stability and the possibility of divorce, disability, and mortality make many aspects of Social Security a gamble. Planning can, therefore, be complex.

Caution must be exercised. A qualifying person (generally one who is married, or was married to an annuitant for at least 10 years) whose personal benefit is less, receives a Social Security benefit equal to 50 percent of that spouse's (or ex-spouse's) benefit [7].

Further caution is required when one spouse has paid the maximum OASDI (\$97,500 x 12.4% for 2007) each year. If the couple takes steps to increase the average earnings for the lower-income spouse, there will be additional OASDI tax due. Depending on many factors, the benefit increase might or might not be worth the additional tax.

Conclusions

Married workers, and other workers, sometimes face decisions as to how earnings affect Social Security benefits. One of these involves the use of family businesses and the partnership/joint venture. This study represents the first stages of research to assist those who stand to benefit, or not.

REFERENCES

- [1] Kalfayan, Garo. *Minimizing Social Security Taxes in Spousal Businesses*, The CPA Journal, September 2003.
- [2] Publication 541: Partnerships (2/2006)
(<http://www.irs.gov/publications/p541/ar02.html#d0e401>).
- [3] *Small Business and Work Opportunity Tax Act of 2007* [H. R. 976, 110th Congress, May 24, 2007]
- [4] Staff of the Joint Committee on Taxation, *Technical Explanation of the "Small Business and Work Opportunity Tax Act of 2007" And Pension Related Provisions Contained in H.R. 2206 as Considered by the House of Representatives*, May 24, 2007 (JCX-29-07)
- [5] Treasury Reg. § 301.7701-2(a).
- [6] Treasury Reg. § 301.7701-3
- [7] www.ssa.org
- [8] 26 C.F.R. § 301.7701-1(a)

HISTORICAL, INVESTORS' AND FINANCIAL ANALYSTS' WEIGHTINGS OF THE CASH FLOW COMPONENTS OF EARNINGS

May H. Lo, Western New England College, (413) 782-1369, mlo@wnec.edu
Emily Le Xu, University of New Hampshire, (603) 862-3318, Emily.Xu@unh.edu

ABSTRACT

This study examines whether or not financial analysts recognize the differential persistence of cash flow components in forming their forecasts of annual earnings. We decompose earnings into four components: accruals, cash distributed to equity holders, cash distributed to debt holders and cash retained by the firm. Consistent with our prediction, we find that financial analysts rank the cash flow components in the same fashion as the historical rankings. Specifically, financial analysts recognize that “cash distributed to equity holders” is more persistent than “cash distributed to creditors” or “cash retained by the firm”. Compared to the historical weightings, financial analysts correctly weight cash retained by the firm. Although analysts appear to underweight cash distributed to equity holders and overweight cash distributed to debt holders, the magnitude of analysts’ mis-weightings is economically small and much lower than the magnitude of investors’ mis-weightings. Moreover, we find that financial analyst coverage has an impact on the weightings of the cash components by both investors and analysts. Our findings complement prior literature and suggest that financial analysts, as information intermediaries, are less biased in processing both the accrual and the cash components of earnings than the investors.

INTRODUCTION

Recently, Dechow, Richardson and Sloan (2006) investigate the persistence and pricing of the *cash* component of earnings, as distinct from the accrual-related anomaly. They decompose the cash component of earnings into three sub-components: (1) cash that is retained by the firm, (2) cash that is distributed to debt holders as a result of debt financing, and (3) cash that is distributed to equity holders as a result of equity financing. They demonstrate that the higher persistence of the cash component (relative to accruals) is entirely attributable to net cash distributions to equity holders. Moreover, investors appear to correctly anticipate the lower persistence of cash distributed to debt holders and the higher persistence of cash distributed to equity holders. However, investors also appear to overestimate the persistence of cash retained by the firm relative to cash distributed to equity holders and debt holders. These results contradict Sloan’s (1996) hypotheses that investors naively fixate on earnings.

To extend this line of inquiry, our study examines whether or not financial analysts recognize the differential persistence of the cash flow components in forming their forecasts of annual earnings. This study focuses on *financial analysts’* (in contrast to investors’) utilization of the information contained in the three cash components because, as information intermediaries, financial analysts play a prominent role in the financial market. Evidence that financial analysts mis-weight the cash flow components in the same manner as do investors is consistent with the

explanation that analysts mislead investors in weighting the cash components. On the other hand, if financial analysts' mis-weightings (if any) of cash components are less than those of investors', this evidence would suggest that analysts are more efficient in processing financial information and that other factors contribute to investors' mis-weightings of the cash flow components. Moreover, Dechow et al. (2006) do not partition their sample according to the quality of information environment, despite recent evidence that richer information environments are associated with more efficient securities pricing. This study examines the impact of the quality of information environment on investors' weightings by grouping the cases based on the level of financial analyst coverage. In summary, the purpose of this paper is to evaluate: (1) whether analysts' weightings of the three cash components are similar in *direction* and *magnitude* to their historical weightings, (2) whether investors' and analysts' weightings (and bias, if any) are similar, and (3) whether the level of sophistication of the information environment affects investors' and financial analysts' weightings of differential cash components.

Our empirical results first show that cash distributed to equity holders is the most persistent among all three cash components, and that investors correctly anticipate the higher (lower) persistence of cash distributed to equity holders (creditors). These results are consistent with Dechow et al (2006). We then show that financial analysts recognize the differential persistence of the cash components. Specifically, financial analysts weight "cash distributed to equity holders" more heavily than "cash distributed to creditors" or "cash retained by the firm", which is consistent with the historical relations. We also find that financial analysts' coverage has an impact on the weightings of the cash components by investors. Consistent with our expectation, the investors' overweighting of accruals, cash retained by the firm, and cash distributed to equity holders are greater for the lower analyst coverage group than for the higher analyst coverage group.

The evidence in this study contributes to the literature in several ways. First, we document that the relative persistence of the three categories of the cash component of earnings remains the same across the two financial analysts' following groups. In both lower and higher analyst coverage groups, we find that the "cash distributed to equity-holders" component is more persistent than the "cash retained by the firm" or "cash distributed to debt-holders" components. This finding suggests that the quality of information environment is not associated with factors that affect the relative persistence of cash flow components, which is consistent with the finding in Elgers et al. (2003) on the relative persistence of the operating cash flows, working capital accrual and other accrual components. Second, we show that investors' overweighting of cash components, as documented by Dechow et al. (2006), is conditioned by the quality of information environment. Using financial analysts' following as proxy for the quality of information environment, we show that investors' overweighting of cash components, except cash distributed to debt holders, is greater for the lower analyst coverage group.

Finally, this study documents the weightings of the three cash components by financial analysts. Prior research has examined both investors' and analysts' weightings of the accrual component of annual earnings. For example, Bradshaw et al. (2001) find that investors overweight working capital accruals, and Elgers et al. (2003) document that financial analysts' overweighting of working capital accruals is much less than the overweighting by investors. In terms of the

weightings of the cash components of annual earnings, Dechow et al. (2006) explore *investors'* weightings and show that investors overweight cash retained by the firm. Our study fills a gap in this line of research by investigating *analysts'* weightings of the cash components. Such an examination allows us to assess whether analysts mislead investors in estimating the persistence of the retained cash component. Our empirical results indicate that financial analysts correctly weight “cash distributed to equity holders” as the most persistent among the three cash components, while investors incorrectly weight “cash retained by the firm” as the most persistent. In addition, our tests of the magnitudes of the weightings show that analysts correctly weight “cash retained by the firm”, while investors over-weight this component. Moreover, although analysts as well as investors overweight “cash distributed to equity-holders” and “cash distributed to debt holders”, analysts’ mis-weightings are significantly less than the over-weightings by investors. Our findings complement prior literature and suggest that financial analysts, as information intermediaries, are less biased in processing both the accrual and the cash components of earnings than are investors.

HYPOTHESES AND RESEARCH DESIGN

Hypothesis Development

To be consistent with Dechow et al. (2006), we decompose annual earnings (E) into total accruals (ACCR), changes in cash (Δ Cash), cash distribution to equity holders (DIST_EQ), and cash distribution to debt holders (DIST_D):

$$E = ACCR + \Delta\text{Cash} + \text{DIST_EQ} + \text{DIST_D} \quad (1)$$

As noted earlier, Dechow et al. (2006) document that the higher persistence of the cash component of earnings is attributable to net cash distributions to equity holders, and that investors correctly anticipate the higher (lower) persistence of cash distributed to equity (debt) holders, although they tend to overestimate the persistence of cash retained by the firms. Prior literature has shown that in some contexts financial analysts, who play a prominent role as information intermediaries in the securities market, utilize financial information better than investors do. This evidence leads us to hypothesize that financial analysts are better able than investors to recognize the differential persistence of various cash flow components in forming their forecasts of annual earnings.

We examine differential persistence from two aspects: relative persistence among the three cash flow components, and absolute persistence of the three cash flow components. With respect to relative persistence, we examine the relative weightings of the three cash flow components that are implicit in analysts’ earnings forecasts. We expect that financial analysts rank the persistence of the cash components in the same order as their historical rankings. With respect to absolute persistence, we focus on the *magnitudes* of financial analysts’ weightings and compare these magnitudes to the historical relations. Specifically, we expect that financial analysts’ weightings of the cash components are not significantly different from their historical weightings. Formally, we state our first two hypotheses as follows:

H1: The rankings of the persistence of the three cash flow components by financial analysts are not significantly different from their historical rankings.

H2: The weightings of the three cash flow components by financial analysts are not significantly different from their historical weightings.

While the focus of this study is on financial analysts, we also investigate investors' weightings of the cash flow components, in order to make complete three-way comparisons among investors' weightings, historical relations, and analysts' weightings. If we find that analysts perform better than investors in weighting the cash flow components, then this evidence complements to the prior literature and suggests that (1) analysts are more efficient than investors in processing not only accruals but also cash flow information, and (2) investors' mis-weightings of cash flow components are not caused entirely by analysts' bias.

Recent evidence indicates that financial analyst coverage is positively associated with the efficiency of investors' information usage. For example, Walther (1997) reports that investors in the securities of lightly followed firms over-rely on time-series predictions of earnings, relative to financial analysts' earnings forecasts. Similarly, Bhattacharya (2001) finds that, for firms with little to moderate analyst following, trading around earnings announcements is more closely related to seasonal random-walk forecast errors than to analysts' forecast errors. Hong, Lim and Stein (2000) find that momentum-based investment strategies are substantially more profitable for firms with low analyst coverage, because the initial under-reaction to value-relevant information, and subsequent returns momentum, are stronger for these firms. Elgers et al. (2001) document delayed security price adjustments to the information in financial analysts' earnings forecasts, and that the magnitudes of the delayed abnormal returns are larger for firms with lower analyst coverage. Overall, these studies provide persuasive evidence that analyst coverage is likely to be an effective partitioning variable, because investors in the securities of highly followed firms appear to use information more efficiently. Accordingly, we adopt analyst coverage as a partitioning device and formulate the following hypothesis:

H3: The level of financial analyst coverage affects investors' and financial analysts' rankings of the persistence of the cash flow components.

H4: The level of financial analyst coverage affects investors' and financial analysts' weightings of the cash flow components.

Overall, we expect that investors' (or analysts') mis-weighting of the cash components of earnings will vary inversely with the level of analyst coverage.

Research Design

In order to examine the historical persistence of earnings components, we estimate the following model:

$$E_{t+1} = \alpha_0 + \alpha_1 ACCR_t + \alpha_2 \Delta CASH_t + \alpha_3 DIST_EQ_t + \alpha_4 DIST_D_t + \mu_t \quad (1)$$

Where E_{t+1} is year t+1 earnings, $ACCR_t$, $\Delta CASH_t$, $DIST_EQ_t$ and $DIST_D_t$ are the accruals, change in cash, cash distributions to equity holders and cash distributions to debt holders components, respectively, of the prior-year's earnings.

In order to examine the weightings of the prior-year earnings components by financial analysts, we estimate the following model:

$$FAF_{t+1} = \beta_0 + \beta_1 ACCR_t + \beta_2 \Delta CASH_t + \beta_3 DIST_EQ_t + \beta_4 DIST_D_t + \mu_t \quad (2)$$

Where FAF_{t+1} is analyst forecasts of annual earnings in year $t+1$. To test H1 and H2, we compare analysts' weightings of the three cash components, estimated coefficients of β_2 , β_3 , and β_4 in expression (2), to their historical persistence, the coefficient estimates of α_2 , α_3 , and α_4 in expression (1), both in direction and in magnitude. Significant results from the equality test of β_2 , β_3 , and β_4 would be consistent with our hypothesis that analysts weigh those cash components differently. We further expect that the weighting for $DIST_EQ_t$ (β_3) would be the highest among the three cash components.

In order to examine investors' weightings of these prior-year earnings components that are implicit in securities prices, we follow the convention to estimate the return/earnings relation using a model that decomposes unexpected earnings into realized earnings and expected earnings. Consistent with the models used in prior studies, we estimate the following model:

$$SAR_{t+1} = \delta(E_{t+1} - \gamma_0 - \gamma_1 ACCR_t - \gamma_2 \Delta CASH_t - \gamma_3 DIST_EQ_t - \gamma_4 DIST_D_t) + \mu_t \quad (3)$$

Where SAR_{t+1} is the size-adjusted security returns in year $t+1$. The investor's weightings of earnings components are measured by γ_i ($i = 1, 2, 3$ and 4). This method was developed in Mishkin's (1985) assessment of rationality in bond pricing and subsequently adapted by Sloan (1996). We estimate this model separately for both the lightly- and closely-followed firms, in order to examine how the quality of information environment (for which analyst coverage is a proxy) affects investors' weightings of the earnings components. To test H3 and H4, we compare the coefficient estimates for the three cash components between these two groups for both investors' and financial analysts' weightings.

RESULTS AND SUMMARY

This study examines whether financial analysts recognize the differential persistence of the cash components, as partitioned by Dechow et al. (2006), in forming their forecasts of annual earnings. The three cash components examined in Dechow et al. (2006) are cash retained by the firm, cash distributed to equity holders and cash distributed to debt holders. Dechow et al. (2006) document that *investors* correctly recognize the differential persistence of cash distributed to equity holders and debt holders while they overestimate the persistence of cash retained in the firm. This study examines how *financial analysts* utilize the information contained in the three cash components in forming their earnings forecasts. In addition, this study partitions the sample based on the quality of information environment (i.e. the level of analyst coverage) and investigates whether the level of sophistication of the information environment affects historical, financial analysts', and investors' weightings of differential cash components.

Consistent with our prediction, we find that financial analysts rank the cash flow components in the same fashion as the historical rankings. Specifically, financial analysts recognize that "cash distributed to equity holders" is more persistent than "cash distributed to creditors" or "cash retained by the firm", which is consistent with the historical persistence rankings. This finding supports our first hypothesis. Compared to the historical weightings, financial analysts correctly

weight cash retained by the firm. Although analysts appear to underweight cash distributed to equity holders and overweight cash distributed to debt holders, the magnitude of analysts' misweightings is economically small and much lower than the magnitude of investors' misweightings. This partially supports our second hypothesis.

Moreover, we find that financial analyst coverage has an impact on the weightings of the cash components by both investors and analysts, supporting our third and fourth hypotheses. Consistent with our expectation, the investors' overweighting on ACCR, Δ Cash, and DIST_EQ are greater for the lower analyst coverage group than for the higher analyst coverage group. However, investors seem to overweight DIST_D for the high analyst coverage group, but not for the low analyst coverage. Finally, we also show that cash distributed to equity holders is most persistent to subsequent year's earnings, regardless of the quality of information environment (i.e. the level of analyst coverage). Overall, our findings document that financial analysts are better able to recognize the differential persistence of the three cash flow components than do investors. These findings complement prior literature and suggest that financial analysts, as information intermediaries, are less biased in processing both the accrual and the cash components of earnings than the investors.

REFERENCES

- Bhattacharya, N., April 2001. Investors' trade size and trading responses around earnings announcements: an empirical investigation. *The Accounting Review* 76 (2), 221-244.
- Bradshaw, M., Richardson, S. and Sloan, R., June 2001. Do analysts and auditors use information in accruals? *Journal of Accounting Research* 39 (1), 45-74.
- Dechow, P, Richardson, S. and Sloan, R., December 2006. The Persistence and Pricing of the Cash Components of Earnings. Working paper. University of Michigan.
- Elgers, P., Lo, M., Pfeiffer, R., and Reece, L., October 2001. Delayed Security Price Adjustments to Financial Analysts' Forecasts of Annual Earnings. *The Accounting Review* 76 (4), 613-632.
- Elgers, P., Lo, M. and Pfeiffer, R., October 2003. Analysts' vs. Investors' Weightings of Accruals in Forecasting Annual Earnings. *Journal of Accounting and Public Policy* 76 (4), 613-632.
- Hong, H., Lim, T. and Stein, J., February 2000. Bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance* 55 (1), 265-295.
- Sloan, R., July 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (3), 289-315.
- Walther, B., Autumn 1997. Investor Sophistication and Market Earnings Expectations. *Journal Of Accounting Research* 35 (2), 157-179.

DELOITTE & TOUCHE USA LLP: TAX SERVICES

INTRODUCTION

For Fiscal Year Ended May 31, 2005, Deloitte Touche Tohmatsu (DTT), a firm that provided accounting services and advice, reported revenues of US \$18.2 billion, up almost 11% from the previous year [Deloitte, 2006C]. William G. Parrett, Chief Executive Officer of Deloitte Touche Tohmatsu, said that sustainable growth was a primary factor in DTT's success in 2005 [Sustaining Growth, 2005]. Sustainable growth was the ability of markets and clients to continue to function and advance as they were faced with serious challenges. DTT faced challenges in maintaining its ongoing commitment to serve clients with professionalism and the best services possible, which played a major role in the company's ongoing success. William Parrett needed to develop an effective differentiating enterprise-wide strategy if Deloitte Touche Tohmatsu was to survive and prosper against aggressive competition over the intermediate and long-term future.

In 2005, DTT had firms located in nearly 150 countries with 120,000 employees worldwide. These firms were called member firms. These firms were legally independent of one another but operated under related names. They provided the functions of DTT all over the world. These member firms were dedicated to providing professional services and advice. Deloitte & Touche USA LLP was one of the many member firms of DTT. In May 2005, Deloitte & Touche USA LLP reorganized. Under the new structure, Deloitte & Touche USA LLP consisted of five subsidiaries, each devoted to a specific function. A subsidiary is a company that is owned or controlled by another company. The subsidiaries of Deloitte & Touche USA LLP were: Deloitte & Touche LLP, Deloitte Consulting LLP, Deloitte Financial Advisory Services LLP, Deloitte Tax LLP, and Deloitte Services LLP.

Deloitte was a part of the professional services industry. The professional services included establishments that specialized in performing activities for others. These activities required a high level of expertise and thorough training. The services provided by Deloitte included AERS (Audit and Enterprise Risk Services), Consulting, FAS (Financial Advisory Services), and Tax. These services were provided through Deloitte's function-specific subsidiaries. Professionals within AERS developed specific knowledge in areas that were significant to clients' businesses. Consulting focused on creating value for clients by providing uniquely integrated capabilities that were supported by the right business structure, style, culture, and size. The consulting and advisory services of Deloitte Consulting LLP were ranked as top-tier, world-class consultancy. FAS offered superlative advice to clients on

essential financial and economic events and transactions. The tax function assisted clients by using the latest tax strategies to significantly impact performance.

As of the end of 2005, Deloitte appeared to be doing well in its industry. Deloitte was consistently increasing revenues and expanding into other countries. The main question to be resolved was how to differentiate Deloitte from its competition and so achieve a winning edge over competitors within intensely competitive, rapidly changing immediate, intermediate, and long-term time frames.

OVERVIEW OF THE INDUSTRY

Tax is a sub-segment of the accounting segment of the professional services industry. Tax services offered varies according to the company providing the services and the services needed by the client. The most successful accounting firms offer a wide variety of services in order to accommodate all clients.

SERVICES

The tax services offered by accounting firms varied between firms. The most common services included international tax, federal tax, state and local tax, and compliance services. Tax professionals were responsible for assisting clients on implementing the advice provided, preparing and reviewing tax returns, and representing clients before federal, state, and local tax authorities.

STAFFING

The staff was responsible for making sure that clients received the appropriate services. In the tax profession, employees entered the company at an entry-level position and moved up through the ranks. A select few had the opportunity to become partner in the firm.

CLIENTS

Tax professionals served a wide variety of clients. The main clients were corporations, nonprofit organizations and partnerships. Corporations were one of the largest sources of revenues for tax. Many of the Fortune 500 companies were corporations, and the main goal of accounting firms was to provide services to these companies. The rules for the taxation of nonprofit organizations were different from for-profit companies. Successful tax professionals were familiar with these rules in order to properly serve nonprofit organizations. Partnerships were also a large source of revenues for tax.

REGULATIONS

There were many laws and regulations applicable to the tax industry. Tax rules and regulations were constantly changing in order to keep up with the changing economy. The main rules and regulations that tax professionals needed to abide by were the Internal Revenue Code, Sarbanes-Oxley, the rules set forth by the Financial Accounting Standards Board, and Generally Accepted Accounting Principles.

MARKETING

Many new clients in the accounting industry were acquired based on the recommendations of existing clients. The Internet was also a main form of advertising for tax services. Firms created websites that explained the services they provided and the industries they provided these services to. Some firms also used sponsorships as a form of advertising and promoting the company name in a positive way.

INDUSTRY OPPORTUNITIES AND THREATS

The following two tables 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
The expansion of clients, creating a need for more tax services	Increased use of outsourcing, decreasing the number of positions available to U.S. tax professionals
Increased use of outsourcing, creating more time for consulting and review	Increasing number of people attempting to prepare their own tax returns
Rules and regulations are becoming easier to follow	Increased competition from other accounting firms

**Table 2
INDUSTRY KEYS TO SUCCESS**

Area	Keys to Success
Services	<ul style="list-style-type: none"> • Show expertise in tax service areas • Services should be high quality • Projects should be completed in a timely and efficient manner • Tax professionals should be aware of new developments in various segments of the industry • Tax professionals should be familiar with rules and regulations of every state and country clients operate in • Help clients comply with the requirements of applicable rules and regulations
Staffing	<ul style="list-style-type: none"> • Have an advanced degree • Receive CPA license • Experience in Accounting • Work well with coworkers • Good client relations, courteous • Always use proper ethics

	<ul style="list-style-type: none"> • Represent clients in front of federal, state and local tax authorities
Clients	<ul style="list-style-type: none"> • Provide services in a wide variety of areas • Must be familiar with all rules and regulations applicable to the client • Minimize taxes and maximize profits • Advise clients of best ways to invest money to receive the best tax benefits
Regulation	<ul style="list-style-type: none"> • Always be up to date on consistently changing regulations • Be familiar with standards of different states • Enforce high ethical standards • Make sure client understands the requirements of the applicable rules and regulations
Marketing	<ul style="list-style-type: none"> • Maintain good relationships with existing clients • Easily accessible, user friendly website • Sponsorships that represent the company in a positive way
Staff Retention	<ul style="list-style-type: none"> • Offer long-term opportunities for employees to develop and grow • Flexible work arrangements • Year-end bonuses and annual raises • Shorter workdays

THE COMPANY

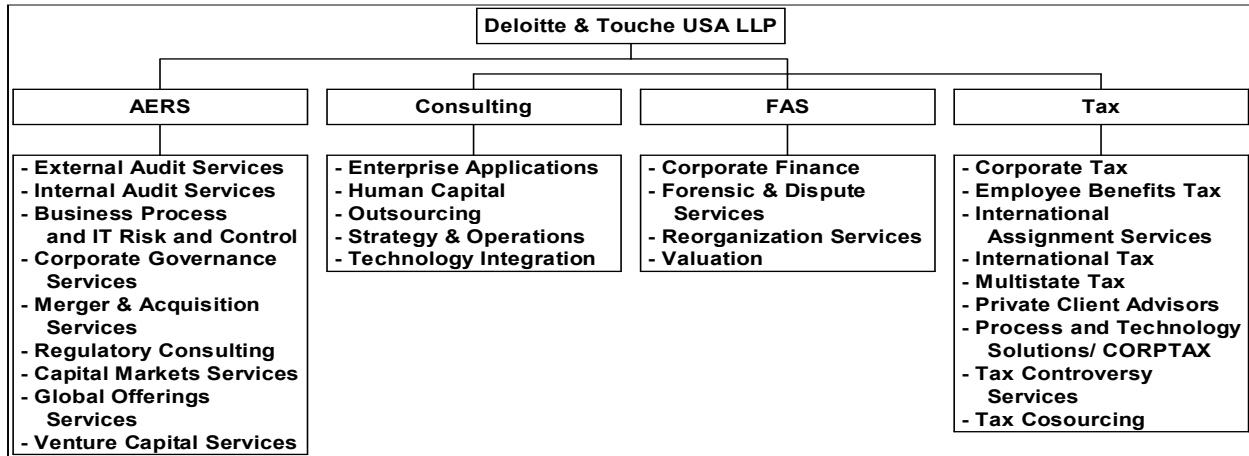
HISTORY

For more than 100 years, clients had relied on Deloitte & Touche USA LLP and its predecessor organizations for a variety of services. In 1893, two accountants, Charles Haskins and Elijah Watt Sells, transformed the way the government did business. They found ways to simplify work and increase efficiency. They became increasingly popular, and in 1895, they opened offices in New York City, Chicago, and London, offering accounting services to the public. In 1898, George A. Touche established a London-based accounting company to assist companies in solving complex business issues. In 1900, Touche joined with John Ballantine Niven and expanded his services to the United States to form the first office of Touche, Niven & Company [Deloitte, 2005B].

In 1947, an accountant from Detroit, George Bailey, launched his own organization. In less than a year, he merged with Touche Niven and A.R. Smart to form Touche, Niven, Bailey & Smart. In 1960, the organization was renamed Touche, Ross, Bailey & Smart, and later in 1969, it became Touche Ross. In the meantime, the company formed by Haskins & Sells was expanding. They merged with 26 domestic organizations and established offices in Canada, Central and South America, Europe, and Japan. In 1990, Deloitte Haskins Sells and Touche Ross merged to form Deloitte & Touche. Deloitte Touche Tohmatsu (DTT) was formed in 1993, when Deloitte & Touche expanded overseas and merged with Tohmatsu & Co, a foreign accounting firm. By 2005, DTT had member firms located in 150 countries around the world, and 120,000 employees [Deloitte, 2005A]. Deloitte & Touche USA LLP was the member firm of DTT located in the United States.

Figure 1 provides a graphical overview of the functions and services provided by the company.

**Figure 1
DELOITTE & TOUCHE USA LLP FUNCTIONS**



Tax

The tax segment is the focus of this study. The tax services segment of Deloitte is referred to as Deloitte Tax LLP, a subsidiary of Deloitte & Touche USA LLP. Deloitte Tax LLP was one of the best companies known for tax advisory. Year after year clients continued to use Deloitte’s services because they were deeply satisfied with the results. Deloitte’s strengths were a result of the high quality services provided by tax professionals. Deloitte tax professionals showed expertise in the tax service area and completed clients’ projects in a timely and efficient manner. Tax services included Corporate Tax, Employee Benefits Tax, International Assignment Services, International Tax, Multistate Tax, Private Client Tax Advisors, Process and Technology Solutions/ CORPTAX, Tax Controversy Services, and Tax Co-sourcing [Deloitte, 2006D, 2006E].

Corporate Tax. Professionals involved in corporate tax services assisted clients with a variety of services including: Accounting for Income Taxes, Accounting Methods and Periods, Global Strategies, Joint Venture and Pass-through Services, Process and Technology, Research and Development, and Strategic Tax Review.

Tax professionals in the Corporate Tax segment helped improve the accuracy and timeliness of clients’ effective tax rate projections and gave the client an idea of how they compared to the competition in their field. Tax professionals helped clients use the most tax-efficient accounting methods, and they applied these methods in order to reduce clients’ taxes. They reviewed clients’ books and records, in order to develop a new tax strategy for improving cash flow. In the global strategies service area, tax professionals ensured that clients were complying with all applicable laws and regulations in each jurisdiction where the company had operations, and they also helped clients strategically structure their companies so they would receive the greatest tax efficiency.

The Joint Venture and Pass-through service area helped clients to structure acquisitions, mergers, dispositions, financing transactions, and day-to-day business transactions in order to generate maximum shareholder value. Deloitte tax professionals helped large enterprises generate maximum shareholder value through the use of partnerships, limited liability companies, business trusts, real estate investment trusts, S corporations, and other pass-through vehicles.

Tax professionals helped clients use technology and associated processes to the greatest advantage. They reviewed companies to ensure that the company was utilizing all the tax deductions and strategies available to them. The review was a thorough process where professionals reviewed the company's tax position to find items that would lower the client's effective tax rate, improve earnings per share, and provide improvements in cash flow by reducing tax payments.

Employee Benefits Tax. Tax professionals within the Employee Benefits segment worked together with the benefit consultants and assurance services to provide clients with the best services. Deloitte tax professionals assisted in the design, implementation, administration, and funding of employee benefits plans, including defined benefit and contribution plans, other post-retirement benefits, workers compensation and health benefit and incentive plans. The key services offered included plan design, plan record keeping and administration, plan qualification, executive compensation and deferred compensation arrangements, excess benefit plans, and merger and acquisition due diligence reviews. The service groups within Employee Benefits Tax included: Business Combination Services, Employee Stock Ownership Plans (ESOP), and Employment Tax.

Tax professionals investigated all areas of employee benefits to ensure that there were no problems with compensation arrangements that the client was not aware of. These problems could potentially expose the company to unanticipated liabilities. These services were also geared towards providing organizations with reasonable financial security strategies that were effective for the company and its executives. Professionals within this segment advised clients and implemented a post-merger integration plan. This plan helped clients bring people, plans, and programs together.

Tax professionals at Deloitte helped companies and business owners with financial liquidity, succession planning, analysis of financial feasibility to achieve business objectives, assistance with structuring of transaction to assure compliance with tax requirements, and transaction implementation and closing assistance. They also offered solutions that helped reduce employment tax liability exposure and created employment tax compliance efficiencies.

International Assignment Services. The International Assignment Services practice assisted clients with addressing and managing the challenges of a global workforce, while adding to the clients' tax compliance, risk management and business performance objectives. The services provided within this practice included: Global Compensation & Benefits, Global Equity Consulting Services, IAS Technology Solutions, IAS-HR Cosourcing, IAS-International HR Consulting, International Immigration Solutions, Law Firm Advisory Services, and Program Optimization Services.

Deloitte tax professionals familiarized themselves with the rules and regulations of issuing stock options, operating employee stock purchase plans, and other types of equity incentive arrangements in every country that the client operated in. This enabled the tax professionals to help the client lower risk and increase benefits for both the company and the employee participants. Employees of Deloitte also helped clients plan, structure, and implement an international program for employees who were working outside of their home countries.

International Tax. The main goal of International Tax was to reduce the client's global effective tax rate. Tax professionals within this practice offered tax-savings ideas that coordinated with the client's overall business objectives and operations. Professionals assisted clients with implementing these ideas and followed up on the situation. The International Tax practice offered the following services: Customs and International Trade Services, Global Earnings Mobility Strategies, International Tax Compliance, Quantitative Consulting Services, and Value Added Tax (VAT).

As a result of Sarbanes-Oxley, reporting requirements were constantly changing, and companies were having a difficult time keeping up with the training and technology necessary to comply with these changes. Deloitte professionals were always aware of new developments in international tax, and they developed solutions for companies that would help them fulfill their requirements.

Multi-state Tax. Many companies operated in a number of states. The tax professionals within the Multi-state Tax practice helped clients manage their state tax liability by using thorough tax approaches that helped companies conduct business and reduce their tax exposure. Services within this practice included: Credits and Incentives, Property Tax, Sales and Use Tax, State Income and Franchise Tax, and State Tax Controversy. Tax professionals familiarized themselves with the rules and regulations in all states the client operated in, and they always kept up-to-date on any changes in the rules and regulations that applied to each client.

Private Client Tax Advisors. Professionals within the Private Client Tax Advisors practice offered planning advice and services to high net worth individuals and executives and owners of closely held businesses. The services offered by the tax professionals within the Private Client Tax Advisors practice covered a broad range including: Business Succession Planning, Comprehensive Financial Planning, Estate, Gift, Trust and Charitable Consulting, Family Office Services, Family Wealth Planning, Individual Income Tax Planning and Compliance, Investment Consulting Services, Life Insurance Consulting, and Partnership and Private Equity Services.

Employees in this service area helped clients determine the best way to distribute their wealth during their lifetime, at death, and after death. Professionals were also able to help partnerships with all aspects of operation beginning with the initial set-up, day-to-day operations, and dissolution.

Process and Technology Solutions/ CORPTAX. Professionals within the Process and Technology Solutions/ CORPTAX practice assisted in implementing process and technology improvements that enhanced efficiency, reallocated resources to value-added activities, and achieved effective integration with CORPTAX software, as well as other information systems. CORPTAX software was software created by employees of Deloitte. It was a web-based system that integrated tax compliance, planning, and reporting. Services within this practice included: CORPTAX, Tax Technology Services, and Tax/ERP Integration Services.

Tax Controversy Services. Deloitte offered a range of tax controversy services that helped resolve difficult IRS matters efficiently. These services included pre-examination audit readiness analysis, examination planning and representation, appeals representation, litigation support, IRS service center matters, IRS penalty and interest computations, and compliance matters involving information reporting and withholding. Tax professionals provided a detailed analysis of taxpayer systems to identify gaps in information gathering and in processes that were necessary for accurately producing information for reporting and withholding compliance. With such detailed analysis, clients were assured that they were abiding by the applicable laws.

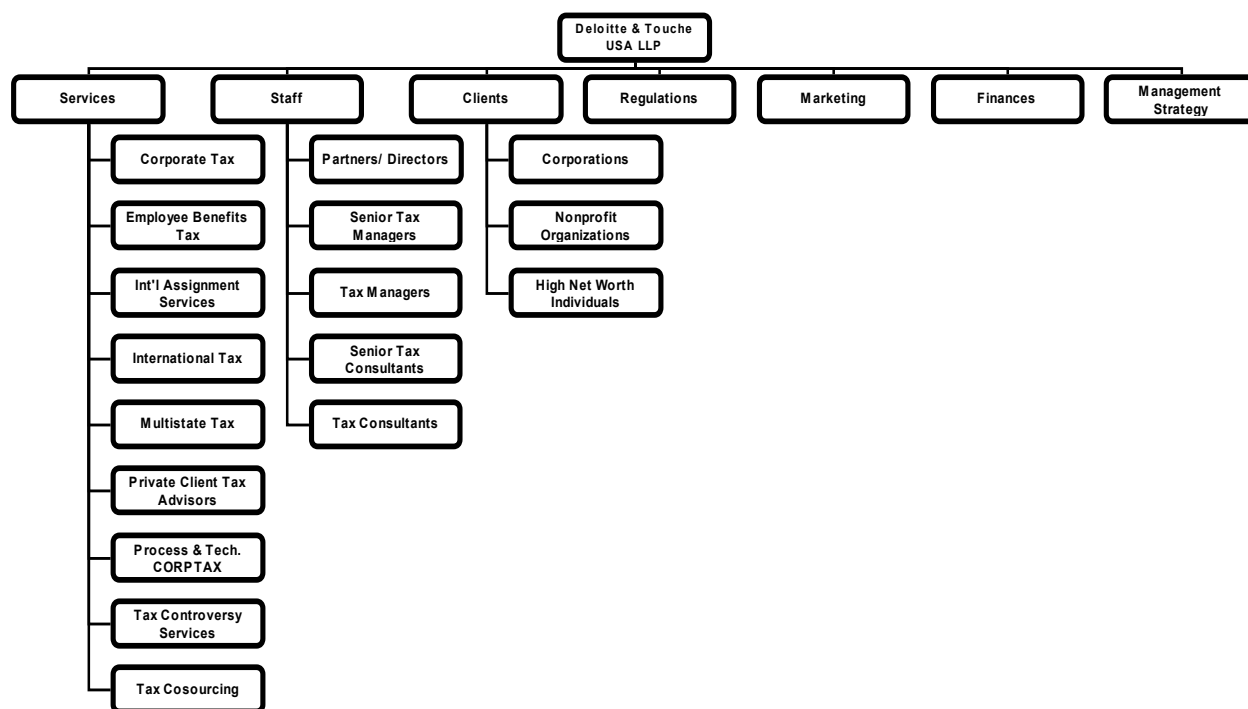
Tax Co-sourcing. Tax Co-sourcing was Deloitte's collaborative approach to outsourcing. Tax professionals in this service area helped clients improve costs and efficiencies, strengthen internal controls, and gain greater access to and control over their organization's compliance activities. Services within this practice included: Business Process Outsourcing, which helped align key supporting functions with business goals and strategies while reducing related costs and risks of outsourced areas; Global Tax Compliance, addressed issues of multinational corporations; Sales & Use Tax Outsourcing, helped reduce the costs of and risks of audit exposure associated with

sales and use tax compliance; and Property Tax Co-sourcing, fulfilled company's property tax compliance obligations in a timely and accurate manner.

Mergers & Acquisitions. When making an acquisition, or going through a merger, companies needed to manage tax risks and ensure that future net cash flows were optimal. Deloitte tax professionals had average expertise in this area. By tax planning at an early stage, clients were able to reduce both the actual transaction tax costs and long-term sustainable tax rate following the transaction. Tax professionals helped clients minimize risks and maximize returns.

Figure 6 provides a breakdown of the Tax segment of Deloitte & Touche USA LLP.

Figure 6
DELOITTE & TOUCHE USA LLP – TAX FUNCTION



STAFF

Deloitte tax professionals were divided among five levels: Partners and Directors, Senior Tax Managers, Tax Managers, Senior Tax Consultants, and Tax Consultants. A weakness of Deloitte was that there was a high turnover rate of the staff. Employees were required to work long hours during tax season, sometimes 65-70 hours a week. Many employees became discouraged as a result of the long hours they had to work. Deloitte did not offer yearly bonuses or other incentives to boost the motivation of their employees. Also, Deloitte did not offer

employees long-term opportunities to develop and grow. As a result of this, employees began to look for other jobs and found positions offering better salaries and less working hours.

A key to success area for staff retention was flexible work arrangements. Deloitte was average in this area in comparison to the competition. Employees were allowed a certain number of personal days, but it was sometimes difficult to take these days if needed during tax season. Women were given maternity leave, and sometimes, upper level staff was allowed to work from home. However, there were still many changes that Deloitte could make in order to provide employees with more flexible work arrangements.

Partners/ Directors

The Partners and Directors of Deloitte were chosen from Senior Tax Managers employed by the company. Partners were employees who purchased shares of the firm. They did not receive set salaries; instead their income depended on the success of the company during the year. Directors had responsibilities similar to those of partner. The difference was that they did not buy shares of the firm. They continued to receive wages from the firm. Partners and Directors were the last employees to look over tax returns and other client information before it was sent out to the IRS or other tax agencies. The Partner or Director was usually the person signing off on the tax information and therefore taking responsibility for the information. Partners and Directors always had to make sure that the information provided to the proper taxing authorities was being reported in an ethical manner. When the partner or director signed off on the client's tax information, they were implying that the information was complete and accurate to the best of their knowledge. Also, it became the partner/director's responsibility to represent the client before federal, state, or local tax authorities if the need arose.

Senior Tax Managers

Senior Tax Managers were either experienced hires from outside the firm, or Tax Managers promoted from within the firm. Senior Tax Managers were required to have at least seven years of general tax experience, experience with consulting and compliance, supervisory skills, and an advanced degree such as a Masters in Taxation was preferred. Managers were also required to have their Certification in Public Accounting. Senior Tax Managers reviewed client information before it was given to the Partner or Director to review. They handled client relations and made sure clients were satisfied with the services they received. It was important for Senior Tax Managers to use proper ethics in all of their decisions.

Tax Managers

The responsibilities of Tax Managers were similar to those of Senior Tax Managers. The main difference between the two levels was seniority and the level of experience. Tax Managers were also either experienced hires from outside the firm or promoted from Senior Tax Consultants from the firm. Tax Managers needed to have at least five years of general tax experience, experience in income tax planning and management, supervisory skills, and an advanced degree was preferred. Tax Managers were also required to have their Certification in Public Accounting.

Senior Tax Consultants

Senior Tax Consultants were either experienced hires from outside the firm or promoted from Tax Consultants within the firm. Senior Tax Consultants were required to have at least two years experience with preparing and reviewing client workpapers and tax returns, they needed to have knowledge of managing projects such as scheduling, budgeting, client correspondence, and billing, as well as research and writing skills.

Tax Consultants

Tax Consultants were usually hired right after graduating from college. Some had only one year of prior experience in the field. Tax Consultants that were hired straight out of college needed to have strong academic credentials, relevant work experience such as interning in the field, a Bachelor or Masters Degree in Accounting or Taxation, and they needed to meet the requirements to sit for the CPA exam before they began working full-time. Tax Consultants that were hired with one or more years of prior experience needed to have experience with general tax, income tax planning, and compliance, a Bachelor's degree in Accounting, Finance or other business related field, and an advanced degree was preferred. It was important at all levels to work well with coworkers, but especially important for Consultants because they interacted with staff at all levels.

Since 2000, Deloitte has been outsourcing tax services to India. Since then, nearly all of the Deloitte Tax U.S. service lines have developed complements in India. The tax professionals in India worked closely with their U.S. colleagues to strengthen client service capabilities. The professionals in India were typically at the consultant level. Staff in India prepared the work and sent it to U.S. professionals for review. Consultants in India were required to have a Bachelor's Degree in an Accounting related field of study, and it was mandatory that they spoke English. Accounting experience at another U.S. multi-national firm was preferred, as well as a Master's Degree. Deloitte employees in India had to be willing to work evening hours, in order to compensate for the time difference

between India and the U.S., because it was essential for them to be in the office the same hours as U.S. professionals.

CLIENTS

The Deloitte tax function served a wide variety of clients. The majority of clients were Corporations, Nonprofit Organizations, Partnerships, and High Net Worth Individuals.

Corporations

Corporations made up a large portion of Deloitte's clients. Deloitte tax professionals became familiar with new tax laws and regulations as they were administered and put into effect. This enabled Deloitte to provide its clients with the most beneficial way of accounting for their money. This usually was whichever method minimized taxes and maximized profits. Deloitte tax professionals' strengths were that they offered services in a wide variety of areas. They were familiar with international tax, federal tax, and state, and local tax. This enabled them to advise their clients in all aspects of their operations.

Nonprofit Organizations

The rules for taxation of nonprofit organizations were different from the taxation of normal organizations because nonprofit organizations did not operate to generate a profit. However, they did hold and disburse money and other things of value. Deloitte's tax professionals possessed strengths in this area by providing services related to the client's needs, and familiarizing themselves with the appropriate rules and advising their clients of the best way to run the organization.

Partnerships

Deloitte tax professionals provided services to many companies that operated as partnerships. Tax professionals were strong in the keys to success areas for partnerships. Deloitte tax professionals were able to provide services in a wide variety of areas, and they applied these services in such a way that would help clients minimize taxes and maximize profits.

High Net Worth Individuals

The individuals that used Deloitte's services usually had a high net value. Deloitte was experienced in a wide variety of services. The individuals that Deloitte provided services to knew that tax professionals would advise them of the best way to divide their money. For certain clients, Deloitte professionals advised that the client

invested their money in trusts or give it to their children as gifts. It was Deloitte's job to ensure clients that they were paying the least amount of tax possible.

REGULATIONS

Each year, the tax services function of Deloitte, issued a Tax Practice Manual. This manual explained the firm's policies, provided background and guidance regarding all applicable regulations and standards, and detailed the proper procedures to comply with the regulations and standards. This helped Deloitte professionals stayed up-to-date on the rules and regulations that they needed to know in order to properly assist clients. The Tax Practice Manual was a clear and concise way for tax professionals to review new rules and regulations. It was well organized and a time saver because professionals just had to look in the manual for the rules instead of other sources which might have taken more time. The manual helped tax professionals convey to clients the requirements of the applicable rules and regulations. Tax professionals always enforced high ethical standards regarding tax regulations.

MARKETING

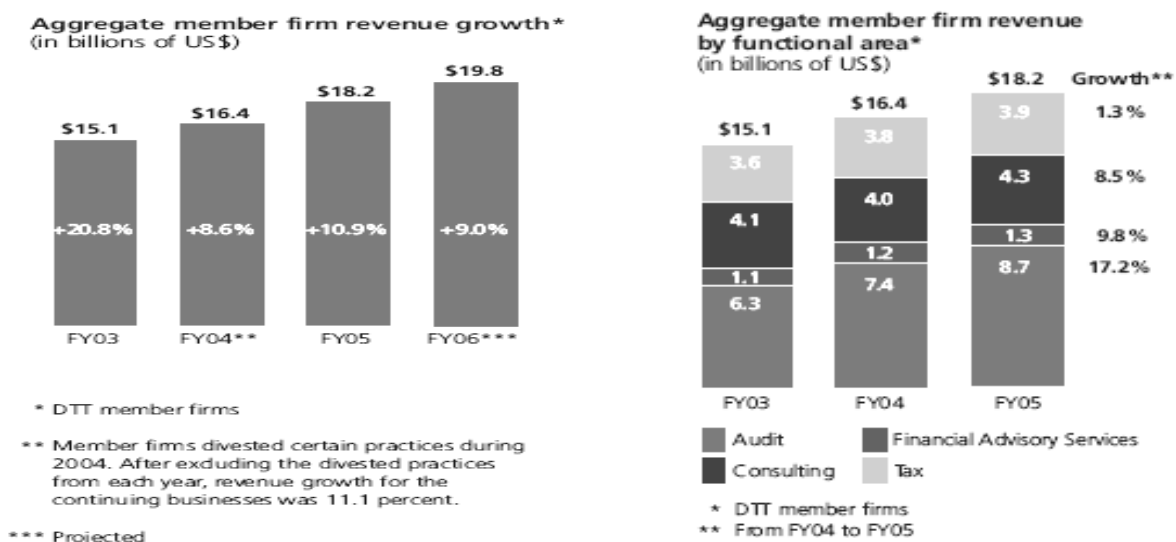
Deloitte did not advertise in the same manner as other companies. Deloitte was one of the Big Four accounting firms and they used sophisticated means of advertising. The company had a very extensive website where a potential client could go and get a detailed explanation of the various services Deloitte offered. The website was easily accessible and user friendly. Deloitte also advertised through sponsorships, donations, and volunteer services. By using these methods of advertisement, Deloitte was able to help others while promoting the company. This made potential clients want to use Deloitte's services because sponsorships and donations made it seem like a reliable and honest company. Another strength of Deloitte tax professionals was to maintain good relationships with existing clients in hopes of getting recommendations from these clients to potential new clients. Deloitte also had its employees volunteer their services on a regular basis. For example, one year a group of employees went to a park to clean and plant trees. This form of promotion put the Deloitte company name into the world for others to see.

FINANCES

Deloitte & Touche USA LLP reported revenues of US\$ 7.81 billion for fiscal year 2005 [Deloitte, 2006A]. This accounted for about 43% of DTT's total revenues of US\$18.2 billion for fiscal year 2005 [Sustaining Growth, 2005]. Strong financial performance enabled Deloitte to provide high-quality services to its clients. In 2005, the demand for Deloitte services, in all four major functions, Audit, Consulting, Financial Advisory Services, and Tax,

remained strong for the third consecutive year. Figure 2 shows Deloitte’s revenue growth of the company as a whole and broken down according to the four major functions.

Figure 2
DELOITTE REVENUE GROWTH



Source: Sustaining Growth (2005). “Deloitte Touche Tohmatsu Worldwide Member Firms 2005 Review”. [Online]. <http://www.deloitte.com/dtt/article/0,1002,sid%253D95275%2526cid%253D103316,00.html>. Accessed February 4, 2006.

Since 2003, Deloitte has been continuously growing and generating more revenues. Each of Deloitte’s functions contributed in some way to the growth of the company. The Audit function grew by 17.2%, Financial Advisory Services 9.8%, Consulting 8.5%, and Tax 1.3%, between 2004 and 2005.

Audit services had the highest growth percentage and also generated almost half of the firm’s revenues. This was due to Deloitte’s commitment to assisting their clients in developing and implementing the appropriate controls that were needed to protect the client’s operations and ensure that they were meeting the regulatory requirements. Tax services had the lowest growth percentage. This might have been a reflection of a weaker market for certain discretionary non-compliance services. After a decrease between 2003 and 2004, consulting services showed a growth between 2004 and 2005. The growth of the Consulting function and Financial Advisory Services could be attributed to the strong global economy and the associated increase in business activity.

Deloitte representatives believed that strategic investments contributed to their strong operating activity results. Each member firm of Deloitte Touche Tohmatsu was continuously investing to enhance its ability to provide clients with the best services possible. Their investments included learning and growth programs to further

develop and enhance professional skills, and developing and improving service methodologies and processes to meet professional, governmental, and industry regulations and standards.

MANAGEMENT STRATEGY

Chet Wood was the CEO and Chairman of Deloitte Tax LLP. In 2005, he was named one of the 50 Most Influential Individuals in Tax, by *Tax Business* magazine. *Tax Business* magazine was the first magazine that focused purely on the tax departments within professional services and law firms, it was written by independent journalists who ensured that readers received unbiased information [Deloitte, 2006B].

In order to succeed in his service area, Chet Wood turned to outsourcing to succeed over the competition. Deloitte tax employees in the U.S. began working with tax professionals in India in order to meet the needs of U.S. clients. The use of professionals in India improved services throughout Deloitte. Professionals in India began doing all the administrative work, such as tracking and billing. This left more time for U.S. professionals to spend consulting with clients. This provided better service, better value, and led to new clients and projects. The team in India provided more consistency, faster turnaround, higher quality, and lower risk.

A major benefit of using professionals in India was that it was profitable for the firm. Deloitte was able to hire professionals in India for salaries much lower than what they would have had to pay professionals in the U.S. This helped the company keep an edge over competitors.

COMPETITIVE ENVIRONMENT

The tax services segment of accounting firms was an extremely competitive area. The accounting industry consisted of large firms that offered similar services and targeted the same clients. The main competitors in the tax segment of the accounting industry were Ernst & Young, Pricewaterhouse Coopers, and KPMG.

Regarding services, Deloitte is equal to or slightly ahead of the competition. The Big Four accounting firms all show expertise in tax service areas, and provide high quality services. Ernst & Young is weak when it comes to international services. KPMG possesses weaknesses in regards to compliance services. Pricewaterhouse Coopers appears to be Deloitte's biggest competitor because they possess many of the same strengths as Deloitte. Deloitte is somewhat behind the competition when it comes to staying up to date on changing state regulations. However, Deloitte is ahead of the competition when it comes to applying rules so that clients reduce their taxes.

Deloitte is on par with the competition in regards to their staff. Each of the Big Four accounting firms strives to hire employees with an advanced degree, CPA license and experience in accounting. Deloitte tax

professionals usually have better relationships with clients in comparison to the competition. The staff at KPMG has a history of not using proper ethics when it comes to making business decisions. This is an area that Deloitte, E&Y, and PwC are all strong in. Deloitte lags behind the competition in the ability to retain staff. This is an area that PwC also possesses weaknesses in. Deloitte does not offer employees enough opportunities for growth and development. Deloitte also requires employees to work too many hours and does not provide flexible work arrangements or year-end bonuses and annual raises.

Regarding clients, Deloitte is average when compared to the competition. When providing services to corporations, Deloitte tends to be ahead of the competition. This is due to the wide variety of services offered and its ability to advise clients how to minimize taxes and maximize profits. When providing services to nonprofit organizations, Deloitte lags behind the competition. Tax professionals at Deloitte are not familiar enough with the services related to the clients' needs, and they are not well familiarized with the rules and regulations applicable to organizations. All Big Four accounting firms are on the same level in providing services to partnerships, and Deloitte is ahead of the competition regarding high net worth individuals.

Deloitte, E&Y and PwC are equally strong in regards to helping clients abide by the applicable regulations and enforcing high ethical standards. All three companies make sure to stay up to date on consistently changing regulations. They are familiar with the standards of different states, and they enforce high ethical standards. These companies also make sure that the clients understand the requirements of the applicable rules and regulations. KPMG lags behind the other Big Four firms in this area. Partners at KPMG have been accused of embezzlement and fraud. This leads the company to be weak in this area. KPMG employees do not enforce high ethical standards, and they do not convey the requirements of applicable rules and regulations to the client.

Deloitte is equal to the competition in the marketing area. The Big Four accounting firms rely heavily on client recommendations and the Internet for advertisement. All four accounting firms have an easily accessible and user-friendly website. In some situations, Deloitte and PwC offer sponsorships that represent the company in a positive way. Deloitte, E&Y and PwC are strong in maintaining good relationships with clients. These companies rely on these relationships to get them recommendations to potential new clients.

LOOKING TOWARDS THE FUTURE

In 2006, Chet Wood, CEO of Deloitte Tax LLP, assured tax professionals and clients of the firm that the company was performing better than expected. However, he still needed to develop a plan to maintain and improve

the quality of services in the future. Chet Wood was faced with two alternatives regarding his staff. He needed to decide which alternative would be most beneficial for his company and produce the most revenues.

The *first alternative* was for the company to stop outsourcing work to India and focus on a well-trained and experienced staff within the U.S. Deloitte would concentrate on hiring the most qualified employees who showed the most potential.

The **benefit** of this alternative was that there were more positions for U.S. citizens. This would decrease unemployment and help the economy. Also, U.S. clients tended to feel more comfortable dealing with U.S. employees. A strength of Deloitte tax professionals was having good client relations, and always showing courtesy to clients. Not all clients wanted their work prepared in India, so hiring additional U.S. employees would create more availability for Deloitte professionals to assist clients.

This alternative was **feasible** because Deloitte operated successfully for many years without outsourcing to India, and they could continue to do so in the future. Deloitte was one of the top four accounting firms long before it began outsourcing to India. If Deloitte hired the right tax professionals within the U.S. it would continue to be part of the top four accounting firms.

This alternative could **win against the competition** because clients would be comforted to know that U.S. professionals were working for them, and would choose Deloitte over other firms that outsourced their work. Deloitte professionals possessed strengths in using proper ethics and worked well with coworkers. By hiring only U.S. professionals, clients would have an opportunity to meet with the tax professionals handling their work. This would give tax professionals the opportunity to build a trusting relationship with clients, and encourage clients to use Deloitte's services in the future. Clients would be able to see for themselves that Deloitte tax professionals had experience in accounting, a CPA license, and that the majority had an advanced degree such as a Masters in Taxation.

The **drawback** to this alternative was that it would be more expensive for the company to hire U.S. professionals to replace the professionals in India. The professionals in India worked for lower salaries than U.S. employees. The U.S. had higher costs of living than India; therefore U.S. tax professionals had to receive higher salaries. This added expense would hinder Deloitte's revenues. **The way around this drawback** would be to increase costs to clients and lower salaries of employees.

The *second alternative* was for Deloitte to improve and expand on the use of outsourcing to India. They would hire well-trained and experienced professionals both within the U.S. and in India.

The **benefit** to this alternative was that with the additional employees, Deloitte would be able to take on more clients and improve relations with current clients. With professionals in other countries preparing the work, U.S. professionals would have more free time to advise clients and more thoroughly review the work. Deloitte professionals would have time to build a strong relationship with clients. With the time saved on the preparation of the client's work, tax professionals could focus on consulting with clients and devising the best way for them to invest their money so that they would pay the least amount of taxes.

This alternative was **feasible** because professionals in other countries required lower salaries, so Deloitte would be able to afford hiring additional professionals. Once the expansion was completed, Deloitte's revenues would increase and cover any expenses that they spent on the expansion.

This alternative would **win against the competition** because Deloitte would be able to accommodate more clients and gain an edge over their competitors. Deloitte professionals would have more time available to them, and with this time they would be able to expand the already wide variety of services offered. Deloitte would be able to branch out and provide services to a larger variety of clients, such as low to high net worth individuals as opposed to just high net worth individuals. It would be too expensive for low net worth individuals to hire Deloitte U.S. tax professionals, but since professionals in India were less expensive, Deloitte would be able to accommodate some low net worth individuals. This would increase Deloitte's presence in the accounting industry and give the company an edge over the competition.

The **drawback** to this alternative was that some clients did not want their work outsourced to other countries. They preferred that U.S. employees prepared their tax information. This might have made some clients uneasy to use Deloitte for their tax preparation services. The **way around this drawback** was to ask clients what their preference was. Deloitte would present clients with an opportunity to choose, and explain to them the benefits of outsourcing, such as decreased costs to the client. Most clients would be likely to choose outsourcing because it would be less expensive for them. For those that preferred not to outsource, Deloitte would still have an experienced staff in the U.S. ready and willing to provide their services to the client.

These are the kinds of decisions that Chet Wood, the CEO of Deloitte Tax LLP is faced with. He had to weigh his alternatives and choose which would be the most efficient and effective alternative for the firm.

REFERENCES

- Deloitte (2005A). "Organization Structure." [Online].
<https://www.deloittenet.com/AboutDeloitte/FirmStructure/DeloitteToucheTohmatsu/default.htm>. Accessed March 1, 2006.
- Deloitte (2005B). "History." [Online]. *http://www.deloitte.com/dtt/section_node/0,1042,sid%253D2273,00.html*. Accessed February 25, 2006.
- Deloitte (2006A). "The Deloitte U.S. Firms: Know the organizational structure and how to communicate it." [Online]. *<https://www.deloittenet.com/AboutDeloitte/FirmStructure/ReorgCentral/default.htm>*. Accessed March 2, 2006.
- Deloitte (2006B). "Chet Wood Named One of the '50 Most Influential Individuals in Tax'." [Online].
<http://www.deloitte.com/dtt/article/0,1002,sid%253D2251%2526cid%253D102177,00.html>. Accessed March 2, 2006.
- Deloitte (2006C). "Facts & Figures – Deloitte: By the numbers." [Online].
http://www.deloitte.com/dtt/section_node/0,1042,sid%253D2282,00.html. Accessed February 4, 2006.
- Deloitte (2006D). "Tax Services." [Online].
http://www.deloitte.com/dtt/section_node/0,1042,sid%253D2152,00.html. Accessed March 1, 2006.
- Deloitte (2006E). "Tax Practice Manual." [Online].
<http://www.deloitte.com/dtt/alert/0,1001,sid%253D2152%2526cid%253D102067,00.html>. Accessed February 12, 2006.
- Sustaining Growth (2005). "Deloitte Touche Tohmatsu Worldwide Member Firms 2005 Review". [Online].
<http://www.deloitte.com/dtt/article/0,1002,sid%253D95275%2526cid%253D103316,00.html>. Accessed February 4, 2006.

THE IMPACT OF CURRENT AND FUTURE RELATIONSHIP VALUES ON RELATIONSHIP QUALITY

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

Liang-Chin Chen, National Yunlin University of Science and Technology, National Kaohsiung
Hospitality College, Taiwan, chin@mail.nkhc.edu.tw

Chiao-I Chen, National Yunlin University of Science and Technology, Taiwan,
g9422711@yuntech.edu.tw

ABSTRACT

The objective of this study is to examine the impact of current and future customer relationship value on relationship quality, incorporating the phase of relationship development as a moderating variable. The relationship value is also divided into economic and social forms for both current and future concepts. University students were recruited as research subject from a Taiwanese university via convenient sampling. Total 504 valid questionnaires were collected. The telephone companies were chosen as research object due to its frequent usage by the university student. The results indicated that current and future relationship values had significantly and positively affected on relationship quality. Furthermore, future social and current economic relationship values have higher impact than that of current social and future economic relationship values. In addition, for economic relationship value, current consideration had higher impact on relationship quality than that of future one. But for social relationship value, future consideration had higher impact on relationship quality than that of current one. Meanwhile, there had interaction effect between current and future social relationship value on relationship quality, but not for economic relationship value. Finally, we also found that the relationship development phase had a moderating effect on the relationship between current and future relationship value and relationship quality.

Keywords: Current and Future Relationship Values, Relationship Quality, Relationship Development Phase

INTRODUCTION

There has been extensive study of relationship marketing ever since the key mediating variables, commitment and trust, were put forward by Morgan and Hunt (1994). Relationship marketing has even become one of the mainstreams in the marketing circle of the 21st century (Barnes 2001 ; Palmatier, Dant, Grewal, and Evans 2006). This may be due to the fact that previous marketing practices was more focusing on the attraction of new customers, particularly while the market is witnessing fast growth. However, strategies on maintaining old customers would become more important with the increasingly saturated market and intensive competition. Therefore, relationship marketing is the critical link for maintaining long term relationships and partnership with old customers. Previous studies on relationship marketing were frequently seen in the B2B market context, which has generated some preliminary satisfactory results (Dwyer, Schurr, and Oh 1987 ; Wilson 1995). For example, the characteristics and mechanisms of relationship between the buyer and seller as well as its impact on performance have been examined. Among those literatures, the relationship quality has attracted most attention from researchers.

In addition, with the redefinition of the marketing by American Marketing Association in 2004, the creation, delivery and communication of customer value has become a core concept of marketing. For this reason, topics on relationship marketing have been set as key research priority for marketing by Marketing Science Institute in the past few years. Moreover, researcher has started to examine the relationship between the relationship value and relationship quality, and still focusing on B2B only (Ulaga and Eggert 2006). Therefore, this study was motivated to focus on the B2C's context in studying the relationship between the relationship value and quality so as to make up B2C's deficiencies in the topic of relationship marketing.

From the theory of cognitive inference, customers may prefer to infer the future consumption decision based on the actual buyer-supplier relationship value (Jacobson & Obermiller 1990). Some studies assume that customers would take the future benefits into consideration (Winter, 1986; Jacobsen & Obermiller, 1990; Boone, et al, 2001). Furthermore, they would also take into account their expectations in addition to the previous and current experiences when reserving such decisions (Lemon et al, 2002). Nevertheless, previous studies concerning the relationship quality may normally neglect the current or future values or interrelation between both two values. Therefore, this study attempts to firstly focus on the issue, from dynamic perspective, to examine whether current and future values may have different impacts on relationship quality. In addition, following the literature of typology on relationship value, it may probably be originated from the economic and social value (Gwinner et al., 1998 ; Reynolds and Beatty, 1999). Then, is there any different impact on relationship quality when customers' purchasing motive is based on these two different value forms considering current and future perspective?

Furthermore, as the development of relationship between the customers and manufacturers may last for a relatively long period of time, customer relationship can actually be divided into different stages of a lifecycle. It is predicable that approaches taken by the subjects in such a relationship may vary at different phases of relationship development (Fournier et al. 2005). However, previous studies may probably neglect such variations at different stages of relationship development when examining the relationship between the relationship value and quality. Therefore, the fourth research question would like to see the moderating effect of the relationship development phase on the relationship between current and future relationship value and relationship quality.

Therefore, the objective of this study is to examine the impact of current and future customer relationship value on relationship quality from the dynamic point of view, incorporating the phase of relationship development as a moderating variable.

LITERATURE REVIEW AND HYPOTHESES

Impact of Relationship Value on Relationship Quality

According to some classical studies, relationship value is an important decisive factor for relationship commitment. The higher the relationship value as recognized by customers is, the higher of relationship commitment would be (Morgan & Hunt, 1994). Relationship value of customers based on the social demands and other perquisites has significant impact on the relationship quality (Hennig-Thurau, 2002). In other words, the higher the relationship value of customers, the higher the relationship commitment would be. This is also favorable for the improvement of relationship quality. Therefore, both parties in the relationship between the customers and enterprises can maintain the customers and, in turn, obtain great profits from the improvement of relationship quality (Zeithaml & Bitner, 2000 ; Hennig-

Thurau,2002). The reason why a customer still has good faith in the original enterprise even if he or she has other choices is that this enterprise can provide a higher relationship value than that provided by its rivals (Ulaga & Eggert, 2006). For a customer, long-term relationship with enterprise is normally established on the relationship value. Furthermore, economic relationship value normally excels the social one (Gwinner, et al. 1998). In other words, such social relationship values as trust, interpersonal relationship and emotion obtained are regarded by customers as the relationship values excelling such economic relationship values as discount from the price.

According to the cognitive inference theory, profits as produced by the price as expected by the customer would affect the result of his or her decision-making (Jacobson & Obermiller, 1990). These cognitive inferences originate from the direct experiences in objects or other relevant information, namely, the relationship value derived from the inference of future price based on the current profits obtained from the existing brands (Winer, 1986). In addition to the satisfaction obtained from the past and current experiences, customers would also take into consideration such factors their expectation for future use and prevention of regret (Lemon, et al. 2002).

H1 : The consideration of current relationship value (economic or social) will have positive impact on relationship quality.

H2 : The consideration of future relationship value (economic or social) will have positive impact on relationship quality.

H3 : Current relationship value will has higher impact on relationship quality than that of future relationship value.

Viewing from relevant literature, it can be discovered that status of relationship value subjected to limited study previously, for example, the current or future relationship value or the concurrent consideration of both current and future relationship values. In recent years, more and more studies have revealed that customers would take into consideration their future profits when making purchase decision (Winter 1985; Jacobsen & Obermiller 1990; Boone, Lemon, & Staelin 2001; Schmitt & Geus, 2006). From the angle of cognitive inference, customers prefer to establish a relationship with an enterprise based on the current relationship value. If the appreciation of future relationship value can be predicated by means of cognitive inference, customers would expect more from the enterprises.

If considering the current and future perspective individually, how will be the different impact of economic and social relationship value on relationship quality? As we discussed above, economic value is more short term, but social value needs to take longer period to build. Therefore, we predict that for those customers who are more considering current perspective will more rely on economic value than social value. If they are more considering future perspective, social value probably will play more critical role than economic value on building customer relationship. Based on above discussions, the hypotheses are established as followings:

H4a : For economic relationship value, current relationship value will has higher impact on relationship quality than that of future relationship value.

H4b : For social relationship value, future relationship value will has higher impact on relationship quality than that of current relationship value.

H5 : There will have interaction effect between current and future social relationship value on relationship quality, but not for economic relationship value.

Moderating Effect of Relationship Development Phase

The phase of relationship development can be described as the marital relationship between the purchasers and vendors. Both parties in the engagement would get married (commitment) in the church after experiencing such procedures as introduction (appeal), acquaintanceship (exploration) and amateness (extension). Nevertheless, separation or divorce (termination) would take place at any time during the development course (Dwyer et al. 1987; Jap & Ganesan, 2000). Therefore, the relationship development phase in this study would like to be divided into 4 major stages. Firstly, the exploration stage: at this stage, both parties would probe into the cost and profits that may be produced by the relationship development. When the expectation is higher than the minimum access preset by individuals, relationship development would become possible. As this relationship at the stage of exploration is not so solid, mutual communication and exhibition of subjects in it would become the major task in developing this relationship (Second & Backman, 1974). Secondly, the build-up stage: with continuous enhancement of interaction between customer and company, social actions would come into being, which would further promote the existence of this relationship via the establishment of relevant regulations and common values. Both parties would establish a relationship of mutual trust and dependence at this stage and would bear more risks that might be produced by the input in developing this relationship. Meanwhile, both parties would voluntarily refrain from seeking other manufacturers for substitution (Frazier, 1983). Thirdly, the commitment stage: both parties would prefer to establish a regular mode through interaction, so as to facilitate the reduction of the uncertainty as well as the increase of profits resulting from the improvement of working efficiency. Furthermore, both parties would have confidence in each other (Dwyer, et al. 1987). Finally, the decline stage: at this final stage, relationship would come to an end in its development course or be in a state of lacking commitment, even if it has not as yet been established (Dwyer, et al. 1987). Furthermore, at least one party in this relationship would have a feeling of dissatisfaction or look forward to more choices for cooperation. Therefore, we expect that the relationship between current and future relationship value and relationship quality will be different at different relationship development phase. The following hypothesis is developed:

H6 : The relationship development phase will have a moderating effect on the relationship between current and future relationship value and relationship quality.

RESEARCH DESIGN

Regarding the operational definition for all variables, current relationship value was defined as any profit obtained by customers in their previous and existing experiences, but future relationship value was defined as any positive utility as expected by customers from the future products or services. Economic relationship value was related to favorable price, economic and favorable services and appreciation, but social relationship value was related to enthusiastic and considerable services, which can facilitate the maintenance of sound social relations or the promotion of adaptation to the cycle formed by relatives and relevant groups. Relationship quality is evaluated in terms of trust, commitment and satisfaction. The aforesaid questionnaire adopts the 7-point measuring indicator as proposed by Likert, in which, any question is provided with written annotation as per the characteristics of the telecommunication industry. The course of relationship development represents the development course of relationship exchange, covering such stages as exploration, extension, commitment and recession. University students were recruited as research subject from a Taiwanese university via convenient sampling. Total 504 valid questionnaires were collected.

ANALYSIS RESULTS

Impact of Current and Future Relationship Value on Relationship Quality

Viewing from the results of regression analysis, the coefficient of current economic relationship value is 0.385 ($t=10.076$, $p<.001$), which indicates that current economic relationship value has significant positive correlation with relationship quality. On the other hand, the current social relationship value is also significantly correlated with relationship quality with coefficient of 0.394 ($t=10.289$, $p<.001$). The same results also happen for both current economic and social relationship values across all individual dimensions of relationship quality such as trust ($b=.228$, $p<.001$ and $b=.346$, $p<.001$ respectively), commitment ($b=.337$, $p<.001$ and $b=.319$, $p<.001$ respectively), and satisfaction ($b=.442$, $p<.001$ and $b=.356$, $p<.001$ respectively). Therefore, H1 was supported. Regarding the future relationship value, the results of regression analysis also indicates that both current and future relationship values have significant impact on relationship quality ($b=0.233$, $P<0.001$; $b=0.495$, $P<0.001$ respectively) and its individual dimensions, trust ($b=.206$, $p<.001$ and $b=.332$, $p<.001$ respectively), commitment ($b=.140$, $p<.01$ and $b=.469$, $p<.001$ respectively), and satisfaction ($b=.268$, $p<.001$ and $b=.481$, $p<.001$ respectively). This is consistent with the expectation of H2. In addition, in order to examine the different impact of current and future relationship value on relationship quality, the regression analysis was utilized to test the hypothesis. The result indicates that both two coefficients are statistically significant in each regression across all dimensions ($p<.001$) and the coefficients of current relationship value are higher than that of future relationship (RQ: $b=0.440$ vs. $b=0.327$; Trust: $b=0.314$ vs. $b=0.251$; Commitment: $b=0.381$ vs. $b=0.262$; Satisfaction: $b=0.450$ vs. $b=0.340$). Therefore, we can conclude that current relationship value will has higher impact on relationship quality than that of future relationship value and the H3 was supported.

Next, we would like to further examine the current and future effects on relationship quality when dividing the relationship value into economic and social values respectively. For economic relationship value, most of the coefficients of current relationship value in each regression for relationship quality and its individual dimensions are higher than that of future relationship value except trust (e.g. $b=.371$ vs. $b=.243$ for relationship quality). In the mean time, the results also indicate that there is no interaction between current and future economic relationship value across all variables. However, for social relationship value, all of the coefficients of future relationship value in each regression demonstrate that they have higher impact on relationship quality than that of current relationship value (e.g. $b=.694$ vs. $b=.543$ for relationship quality). On the other hand, there also have significant interaction effects between current and future social relationship value on relationship quality ($b=.482$, $p<.01$) and other individual dimensions. Therefore, H4a, H4b, and H5 were all supported.

Analysis of Moderator Variables of Current and Future Relationship Values

In order to examine the moderating effect of relationship development phase on the relationship between relationship value and relationship quality, current and future relationship values were put into regression analysis for each phase of relationship development. The results indicate that although current and future relationship values had significant effects at the exploration stage, the key influence were coming from current and future social relationship value ($b=.253$, $p<.01$ and $b=.218$, $p<.05$), not from economic value ($b=.122$, $p>.05$ and $b=.143$, $p>.05$). At the build-up stage, future social value ($b=.394$, $p<.001$) seems playing more important role than future economic value ($b=.200$, $p<.05$). At the commitment stage, current social value had no significant effect on relationship quality ($b=.031$, $p>.05$) despite overall current value had. At decline stage, although overall future value had no influence, the

future economic value, not future social value, still had positive relationship with relationship quality ($b=.313, p<.05$). Based on above analysis, we found that the relationship development phase have a moderating effect on the relationship between current and future relationship value and relationship quality. Therefore, H6 was supported.

DISCUSSIONS AND CONCLUSIONS

The objective of this study is to examine the impact of current and future customer relationship value on relationship quality, incorporating the phase of relationship development as a moderating variable. The relationship value is also divided into economic and social forms for both current and future concepts. The results indicated that current and future relationship values had significantly and positively affected on relationship quality. Furthermore, the four types of relationship value, combining current and future consideration with economic and social types, all have significant positive relationship with relationship quality and its individual dimensions, but surprisingly, future social and current economic relationship values have higher impact versus current social and future economic relationship values with lower impact. In other words, future social and current economic relationship values are the critical factors for the establishment of relationship quality between customers and enterprises.

In addition, for economic relationship value, current consideration had higher impact on relationship quality than that of future one. But for social relationship value, future consideration had higher impact on relationship quality than that of current one. Meanwhile, there had interaction effect between current and future social relationship value on relationship quality, but not for economic relationship value. However, if considering current and future perspective, future social relationship value had higher impact on relationship quality than that of future economic relationship value, but current economic relationship value did not find higher impact on relationship quality than that of current social relationship value. Therefore, if a brand positions it on economic value, current consideration will be more important than future consideration. On the contrast, future perspective may need to put it into its positioning if it claims on social value. In the meantime, the synergy of current and future social value could be taken into consideration due to its significant interaction effect.

Finally, we also found that the relationship development phase had a moderating effect on the relationship between current and future relationship value and relationship quality. At the exploration and commitment stages, current and future relationship value showed significant influence on relationship quality. However, only future relationship value had significant relationship with relationship quality at the build-up stage, but not for current relationship value. Meanwhile, at the decline stage, the current relationship value played an important role of influencing relationship quality, but the future relationship value has no significant effect on relationship quality. This implies that marketer should plan different relationship marketing program for different customer based on its phase of relationship development.

In summary, this is few study of relationship marketing focusing on B2C context and using dynamic perspective to examine the correlation between relationship value and relationship quality. The current and future consideration, in according with economic and social values, and the phase of relationship development are the two dynamic perspectives reflecting its innovation to the literature of relationship marketing. We believe that the findings of this study will make some contribution to the academic and practical fields of customer relationship management.

Applications of the Habitual Domains Theory: Individual Life and Enterprise Organizations

Ying-Yu Ian Chen, Kainan University, Taiwan
(886)933-903354 , yyianchen@gmail.com

ABSTRACT

This paper intends to discuss the Habitual Domains Theory (HD) conceptualized by Dr. Po-lung Yu from the perspectives of individual life and enterprise organizations.

HD has created many new dimensions in personal creativity and enterprise innovation by drawing research from psychology, cranial nerves science, system science, and business management. Yu pointed out that human behaviors tend to be stabilized and static by following a routine life pattern. As a result, individual thought and perception is confined without flexibility. The best way to break through the static ways of life is to stay alert in one's own body and soul. By so doing, this will allow each individual to examine his/her blind habitual domains on the daily bases, and eventually enhance one's efficiency and effectiveness from life and work.

The application of HD to the enterprise is also valid, thanks to the commonality between individual behaviors and organizational structures. The latter require transparent information for decision-making and action by keeping their open-mindedness and sensitivity for market change.

Examples will prove that HD will shed some lights on individual life as well as enterprise organizations.

Key words:

Habitual Domains (HD). Enterprise Organization Habitual Domains (EOHD)

INTRODUCTION

The collection of thinking, judging, and acting in response to various events and stimuli is called our Habitual Domains (HD). HD may be viewed as our human software, which drives our brain (hardware) to work. However, it has been pointed out and mathematically proven by Po L. Yu that our mental programs which includes thinking, judging, acting, memorizing, and response to various events will gradually become stable and contained in a certain domain(Yu,1995). Therefore, we will develop habitual ways of thinking, judging, acting, and responding.

HD is unwritten, invisible, but always present and affects all our decisions and behavior. Just like a turtle's shell, which follows the turtle; wherever we go, our HDs follow us. Thus, individuals and organizations will develop habitual ways of thinking, judging, acting, and doing business.

APPLICATIONS TO INDIVIDUAL LIFE

HD affects and contains all our decisions and behavior. Without large and flexible HDs, winning strategies for solving significant problems and conflicts and adopting to change will be difficult to achieve. Let us take a look at two examples.

Example 1: The Wisdom of the Retiring Chairman

There was a retiring chairman who wanted to select a successor from his two final candidates (A and B). All three individuals were good at horseback riding. The chairman invited A and B to his farm and gave each finalist an equally good horse. He pointed out the course for the race and the rules, saying "From this point, whoever's horse is slower in reaching the final point will be the new chairman." The rule of the horse racing was outside the habitual ways of thinking of A and B. Both of them were puzzled. If you were the candidate for the job, what would you do?

After a few minutes, A suddenly jumped out of the constraint of his habitual domain. He quickly mounted B's horse and rode as fast as possible, leaving his own horse behind. When B realized what was going on, it was too late. A became the new chairman.(Datta,1991)

Obviously, the problem posed by the chairman in the above example was unique and did not fit the individuals' habitual ways of thinking. Consequently, it generated uncertainty and anxiety among both players until 'A' was able to 'expand' his habitual domain thus effectively 'solve' the problem.

Example 2: Long Hairs & No Hairs

One successful businessman often came home at midnight because of engaging in social activities. So his wife was extremely worried about her husband, that he might be having a love affair outside. When her husband comes home, she would secretly check her husband's clothing. If she found some long hairs that were not hers, she would have a quarrel or fight with her husband. After a long while, she could not find any more long hairs on her husband's clothes.

One day this man came home and saw his wife crying sadly. He curiously asked his wife: “Why do you cry so sadly?” His wife answered: “I have not been able to find any long hairs in your clothes for almost two months!” He said: “Isn’t this just what you want?” His wife stared at him with the coldest eyes and said: “How could you downgrade yourself so low as to love women without hairs?”(Yu &Chiang, 1999)

The above example shows us that “we do not first see, then believe; we believe first and then see.” In other words, people selectively perceive what they expect and want to see. The wife assumes that her husband comes home late because of his love affairs. But what if she changes the assumption to “My husband comes home late because he is working hard”? This example also illustrates that it is not easy to quickly and flexibly change our mental programs.

APPLICATIONS TO ENTERPRISE ORGANIZATIONS

Just like individual, organizations too, overtime, develop their own habitual domains. Here are two examples.

Example3: Dog Food

A dog food company designed a special package that is not only was nutritious, but also could reduce dog’s weight. The statistical testing market was positive. The company started “mass production”. Its dog food supply was far short from meeting the overwhelming demand. Therefore, the company doubled its capacity. To their surprise, after one or two months of excellent sales, the costumers and the wholesalers began to return the dog food package, because the dogs did not like to eat it.

Clearly, a decision trap was committed by using statistics on buyers, not on the final users (dogs). They used statistical method on the “wrong” subject and committed the trap. If They could include the buyers and the users, the decision traps and wrong decisions might be avoided.(Yu & Lai, 2004)

Example4: Successful Consultant

In 1987 Dr. Po L. Yu undertook a consulting engagement with United Telecom dealing

with its corporate strategic positioning. In each of my interviews with the officers (president, executive vice presidents, vice presidents, and so on) He started this approach:

“You have been very successful and have an important position in the corporation. You are very intelligent and knowledgeable, especially about your business of telecommunications. Thus you are qualified to be my teacher. Will you please be so kind to treat me as your student?” The resulting process, which applied the deep-and-down principle, gave me a broad, deep, and detailed knowledge of the company.

Each mentor offered a different personal HD and a different HD for the company. Using projection, association, and alternating principle (as well as other ways of achieving deep knowledge), hard work led to an expanded HD for the firm that integrated those of its officers and staff. This new HD enabled me to see-and suggest to the managers-some new ideas that had formerly been outside their HDs. That experience was a rewarding success which was due in large measure to the leadership and support of Paul Henson, Telecom’s Chairman.(Yu, 1991)

This experience suggests a general idea: Developing a corporate strategic position can be seen as developing the corporate HD in such a way that the new concepts can be comprehensive and flexible enough to include and take advantage of a wide spectrum of possible business opportunities , many of which can be only dimly perceived at a given time. Such an expanded and enriched corporate HD can also help in avoiding business difficulties.

HOW TO EXPAND AND ENRICH HABITUAL DOMAINS

In order to study habitual domain effectively, we distinguish HD by four basic elements:

(I)The potential domain (PD): the collection of ideas and mental programs that can potentially be activated in our mind.

(ii) The actual domain (AD): the set of ideas and mental programs that are actually activated at this particular time and space.

(iii) The activation probabilities (AP): the probabilities that ideas and mental programs in PD also belongs to AD.

(iv) The reachable domain (RD): the set of ideas and mental programs that can be attained from a given set in a HD.(Yu & Chiang, 1999)

Obviously, having larger habitual domains is the key to solving more challenging problems. On the other hand, can we acquire a larger habitual domain? Luckily one of the most important characteristics of habitual domains is that they can be expanded

over time. A habitual domain expands first within the reachable domain, then it expands by jumping out of the reachable domain, and finally it expands by jumping out of the potential domain. One way to expand an individual's habitual domain is through interaction with different habitual domains.(Fang, 2001)

There are three toolboxes which helps us expand and enrich our habitual domains and avoid decision traps. We list them in the following three tables.

Table 1: 9 Principles of Deep Knowledge (Yu & Lai, 2004)

1. Deep and Down Principle
2. Alternating Principle
3. Contrasting and Complementing Principle
4. Revolving and Cycling Principle
5. Inner Connection Principle
6. Changing and Transferring Principle
7. Contradiction Principle
8. Cracking and Ripping Principle
9. Void Principle

Table2: 8 Methods for Expanding Habitual Domains (Yu & Lai, 2004)

1. Learning Actively
2. Take the Higher position
3. Active Association
4. Changing the Relative Parameter
5. Changing the Environment
6. Brainstorming
7. Retreat in Order to Advance
8. Praying or Meditation

Table3: 7 Self-Perpetuating Operations (Yu & Chiang, 1999)

1. Everyone is a Priceless Living Entity
2. Clear, Specific and Challenging Goals Produce Energy for our Lives
3. There are Reasons for Everything that Occurs
4. Every Task is Part of My Life's Mission
5. I Am the Owner of My Living Domain

6. Be Appreciative and Grateful and Don't Forget to Give Back to Society
7. Our Remaining Lifetime is Our Most Valuable Asset.

CONCLUSIONS

The new economic world order in which we live demands individuals and organizations who are flexible, open-minded, and possessing large HDs: HDs those are not rigid, which respond flexibly, as opposed to habitually, to events and stimuli.

Without large and flexible HDs, winning strategies for solving significant problems and conflicts, and adopting to change in this new environment will be difficult if not impossible to achieve.

We thus have to instill in ourselves as well as in our colleagues and co-workers a culture of continually expanding and enriching our HDs.

Reference upon request

The relationship among learning organization, internal marketing and organizational commitment

Yafang, Tsai

Department of Health Policy and Management, Chung-Shan Medical University, Taiwan
+886-4-24730022 ext 17176 avon611@yahoo.com.tw, yafang@csmu.edu.tw

ABSTRACT

This study probes into the relationship of learning organization, internal marketing activities and staffs' organizational commitment via an empirical study in a medical center that has implemented learning organization for several years. We take the nurses who work in the medical center as samples. The results show that the dimensions of experience, structure, culture and information of hospital learning organization have significantly positive correlation to value commitment, work-hard commitment and retaining commitment of staffs' organizational commitment. The vision and training dimension of hospital internal marketing have significantly positive correlation to incentive and value commitment and work-hard commitment of staffs' organizational commitment. And the incentive and communication dimension of internal marketing have significantly positive correlation to value commitment and work-hard commitment dimensions of staffs' organizational commitment.

Keywords: Learning organization, internal marketing, organizational commitment, human resource management, health care management

INTRODUCTION

Because of the trend of globalization, organizations have to face the dynamic competition. Organizations have to continually learning in order to survive, and it makes the managers pay more attention to the promotion of learning organization. Joaquin & Ricardo[1] argue that before you can innovate effectively you have to prepare your organization to be open to innovation by creating a learning organization. In the era of knowledge-driven economy, knowledge and learning is the vigor that can not only push individual, organization, society to progress continuously but also can expand business development. The most important assets of organizations is personnel and knowledge, learning organization emphasize that individual in organization is the focal point that can share organizational information and knowledge creation. Organization not only can raise the creativity and efficiency via individual learning, but also can satisfaction employees' job demands and improve organizational commitment of staffs via promotion of learning organization. Internal marketing is one of ways to transmit the vision and mission to staffs. Managers use internal marketing to educate and improve staffs' ability in order to ensure that staff can accomplish assigned mission, and they can retain staffs via raising the centripetal force by incentive activities of internal marketing.

Research objectives and importance of the investigate

Organizational learning is a management approach applicable to a wide range of organizations, including businesses, government agencies, and schools [2]. Learning is seen as a systems-level phenomenon that stays within the organization regardless of the "players" involved. But there is little research focus on health care institutes. Studies on internal marketing primarily adopt the service industry as the research object, but most researchers still view commercial organizations as the research object. However, administrators of nonprofit organizations, such as hospitals, should pay greater attention to improving internal management process, for example through downsizing, reorganization, and so on, owing to the influence of changes in medical environment and health policy. Cooper & Cronin [3] though focused on nursing home care and their focus was limited to internal marketing activity, but the main goal of internal marketing is to enhance service quality for external customers [4]. According to the service-value chain [5], the goal of internal and external customer satisfaction can best be achieved through the influence of certain mediator variables. That is, internal marketing can improve internal customer satisfaction, which can serve as a mediator to external customer satisfaction [6]. Other variables, such as service process design, have also been found to be mediators of external customer satisfaction. Consequently, this study aims to extend past research by examining whether organizational commitment is correlation with internal marketing.

LITERATURE REVIEW

Learning organization

Definition of a learning organization

Organizational learning is discipline based and analytic, the newer concept of developing a holistic" learning organization " is multidisciplinary and emphasizes action and the creation of an "ideal type" of organization[7]. Learning organizations are able to transform themselves while facing multiple, complex forces. Steven et al. [8] knows if the is indeed possible, then hospitals-which must deal with at least as much constant change, competition, and external pressures as any other organization-would stand to gain from becoming learning organizations. These organizations use generative learning for continuous evaluation and feedback to reprogram how information is processed. Thus they encourage experimentation and system wide thinking.

Features of a learning organization

Senge [9] listed the five disciplines of learning organization, including: personal mastery, mental models, building shared vision, team learning, and systems thinking. Huw et al. [10] propose five key features of a learning organization, they are open systems thinking, improving individual capabilities,

team learning updating mental models, and a cohesive vision. Learning organization actions are purposive behaviors and patterns of interaction in organizational members' approach to work that provide opportunities for learning and the demonstration of shared responsibility and collective competence in addressing organizational goals[2]. There are different labels may be used to describe dimensions that reflect this action component of organizational learning, processes such as team orientation, innovation, involvement, information flow, tolerance for error, and results orientation[11].

Internal marketing

The concept of internal marketing first appeared in the service marketing literature in the early 1980s[12]. Internal marketing stresses that firms can retain employees through motivating, mobilizing, cooperating, and managing, continuing to improve the services provided to external customers [13]. Internal marketing is also focused on training and inspiring employees to serve clients well [14]. Tansuhaj et al. [15] believe that internal marketing stresses various plans that emphasize employee development and that a complete internal marketing program must necessarily include employee recruitment, training, encouragement, communication and retaining activities that lead to changes in employee attitudes and behavior. In other words, internal marketing is a company's human resources management activity for the successful hiring, training and encouraging of employees, and providing better services to clients [16]. George & Grönroos [17] more clearly point out that internal marketing is a philosophy of managing human resources from the marketing perspectives [18].

Organizational commitment

The extensive body of scientific studies on organizational commitment, largely focused on organizational behavior/psychology, has produced various definitions of the factor construct and considerable discussion of its development, consequences, and measurement [19]. At a general level, most researchers agree that organizational commitment describes a psychological state that characterizes the relationship of an employee with the organization for which they work and that has implications for their decision to remain with the organization [20]. Organizational commitment comprises employee attachment to their organization [21] [22]. Meyer and Allen [23] have identified three types of organizational commitment. Affective commitment describes individual emotional attachment, identification with, and involvement in a particular organization. Furthermore, continuance commitment reflects employee awareness of the costs of leaving an organization. Finally, normative commitment reflects individual sense of obligation to remain in an organization.

The relationship among learning organization, internal marketing and organizational commitment

Health care organization is one kind of knowledge-denseness organization, managers can ensure medical quality and technology, and then can meet the requirement of patients for medical service via staffs' continually learning. Here is a question that how to encourage staffs to learn continually? We can make it by molding the learning organization. Hospital administrators not only encourage staffs solve the problems on their job by systematic thinking, but also should urge team learning. Medical staffs can transfer knowledge, bring new concept and upgrade their abilities via communication and coordination of members of medical team. Once staffs' abilities have been upgraded, they can face the challenge of medical environment changing, solve problems on their job easily, and also can raise their commitment to organization.

H1: There is correlation between learning organization and organizational commitment

Internal marketing is based on the notion that organizations providing secure employment are committed to their employees [24]. This commitment should be returned by employees by demonstrating enhanced job satisfaction, loyalty, and trust [25]. Thus, firms can use internal marketing programs to enhance employees' job satisfaction [6]. Internal marketing emphasizes that firms can hire employees suitable for the organizational mission while providing them with a clear vision and goals [4]. Organizational commitment indicates the degree to which individuals are attached to the organization and identify with organizational goals [26]. Internal marketing assumes

that organizations can improve internal customers' job satisfaction through human resource management. Previous studies have shown that employees' organizational commitment is positively associated with job satisfaction [27]. Satisfied employees will tend to adopt positive attitudes toward the service encounter and thus improve service quality for customers. Therefore, internal marketing is positively related to employee organizational commitment [28].

H2: There is correlation between internal marketing and organizational commitment

METHODOLOGY

Sampling

We take all of nurses as samples working in the medical center that has implemented learning organization for several years in middle Taiwan. Because nurses are always busy at their work, we adopt convenience sampling to choose nurses who are willing to be our samples. We send out 100 questionnaires and take effective 45 back.

Data collection

Learning organization. We refer to the concept of learning organization proposed by Hiatt-Michael [11] to design variables. **Internal marketing.** We used the questionnaire developed by Money & Foreman. Money & Foreman [29] adopted a case study to develop internal marketing measuring tools, and this scale has subsequently been widely adopted [30]. **Organizational commitment.** Organizational commitment (OC) was measured by a fifteen-item survey that was developed by Mowday et al. [22]. These works were implemented by a questionnaire survey using the 4 point Likert scale.

RESULTS

Descriptive statistics

The mean value of learning organization lies between 2.42 and 2.97(the scale is 1 to 4). The mean value of internal marketing lies between 2.29 and 2.78. Moreover, the mean value for employee's organizational commitment lies between 2.48 and 2.82.

Inferential statistical analysis

Factor analysis

According to the analysis, The Cronbach's alpha of total learning organization scale is 0.898, the accumulated variation is 68.579%; the Cronbach's alpha of total organizational commitment scale is 0.767, the accumulated variation is 57.898% ; the Cronbach's alpha of total internal marketing scale is 0.930, the accumulated variation is 63.623%. The result of factor analysis show us there are four dimensions of learning organization, we name them by each characteristic as experience, structure, culture and information. There are three dimensions of organizational commitment as value commitment, work-hard commitment and retaining commitment. There are two dimensions of internal marketing as vision and training, motivation and communication.

Correlation analysis

We adopt Pearson's correlation analysis to analyze the relationship between four dimensions of learning organization and three dimensions of organizational commitment, the results are as follows: (1) there is significantly positive correlation between experience dimension of learning organization, value commitment and work-hard commitment of organizational commitment. ($r=0.619$, $p=0.000$; $r=0.414$, $p=0.003$) (2) there is significantly positive correlation between structure dimension of learning organization, value commitment and work-hard commitment of organizational commitment. ($r=0.503$, $p=0.000$; $r=0.422$, $p=0.003$) (3) there is significantly positive correlation between culture dimension of learning organization, value commitment and work-hard commitment of organizational

commitment. ($r=0.575, p=0.000$; $r=0.456, p=0.001$)

The results of correlation analysis between two dimensions of internal marketing (IM) and three dimensions of organizational commitment (OC) are as follows: (1) there is significantly positive correlation between vision and training dimension of IM, value commitment and work-hard commitment of OC. ($r=0.413, p=0.004$; $r=0.339, p=0.018$), (2) there is significantly positive correlation between motivation and communication dimension of IM, value commitment and work-hard commitment of OC. ($r=0.339, p=0.018$; $r=0.663, p=0.000$)

CONCLUSIONS

Because of the violent variation of the industrial environment, we have to improve internal managing abilities and sustain competitive advantage continuously. Learning organization and internal marketing is thought to be a useful way to succeed. The employees of hospitals delivery health care service to patients, the competence and skill of employees will affect the service quality. The administrators of hospitals have to improve employees' competence first, and then can increase operative abilities, and learning organization is one of ways to improve employees' competence. As staff's competences have been improved, administrators should devote to keep and select outstanding employees and avoid the same business recruiting from ours. Internal marketing is one action of human resource management to keep excellent staffs. The results prove that there is positive correlation between learning organization and internal marketing and staffs' organizational commitment. We suggest that hospital administrators can apply these two ways to improve employees' competence and centripetal force, and they can face the challenge of the violent variation.

The main purpose of the study is to develop framework, we only take staffs of one hospital to be our sample and examine the preliminary framework. We will enlarge the sample size to improve the extrapolation of conclusion in the future.

REFERENCES

- [1]Joaquin Alegre & Ricardo Chiva, "Learning to innovate." *MIT Sloan Management Review*, 2007, 49(1), 10-11.
- [2]Bowen, Gary L., Roderick A. Rose, William B. Ware " The Reliability and validity of the school success profile learning organization measure." *Evaluation and Program Planning*, 2006, Vol.29, 97-104.
- [3]Cooper, Jack & John J. Cronin." Internal marketing: A competitive strategy for the long-term care industry." *Journal of Business Research*, 2000, Vol. 48,177-181.
- [4]Reynoso J.F., Moores B. Internal Relationship, in: Buttle F, editor. *Relationship Marketing: Theory and Practice*, London: Paul Chapman Publishing, 1996.
- [5]Heskett, James L., Jones, Thomas O., Loveman, Gary W., Sasser, Earl W. & Schlesinger, Leonard A. " Putting the service-profit chain to work." *Harvard Business Review*,1994, March-April, 166.
- [6]Bansal, Harvir, S, Mendelson, Morris B., Sharma, Basu."The impact of internal marketing activities on external marketing outcomes." *Journal of Quality Management*, 2001, Vol. 6, 61-76.
- [7]Easterby-Smith, M. "Disciplines of organizational learning: Contributions and critiques." *Human Relations*, 1997, 50(9), 1085-1113.
- [8]Steven B. Dowd, Ed D, RT. "Organizational learning and the learning organization in health care." *Hospital Materiel Management Quarterly*, 2000,21(3), 1-3.
- [9]Senge, P.M. *The fifth discipline: the art and practice of the learning organization*. New York: Doubleday. 1990.
- [10]Huw T. O. Davies, Sandra M. Nutley, "Developing learning organizations in the new NHS." *British Medical Journal*, 2000, 8(320), 998-1001.
- [11]Hiatt-Michael, D.B., "Schools as learning communities: a vision for organic school reform." *The School Community Journal*, 2001, Vol.11, 113-127.
- [12]Frost, Frederick A. & Kumar, Mukesh. "Service quality between internal customers and internal

- suppliers in an international airline.” *The International Journal of Quality & Reliability Management*, 2001, 18 (4), 371-386.
- [13] Joseph, Benoy W. ”Internal marketing builds service quality.” *Journal of Health Care Marketing* 1996, 16 (1), 54-59.
- [14] Grönroos, C. Internal marketing-Theory and practice, in: Bloch TM, Upah GD, Zeithaml VA, editors. *Services marketing in a changing environment*. Chicago: American Marketing Association, 1985.
- [15] Tansuhaj, P., Randall D. & McCullough, J. ”Applying the internal marketing concept within large organizations: As applied to a credit union.” *Journal of Professional Services Marketing*, 1991, 6 (2), 193-202.
- [16] Kotler, P. *Marketing Management- Analysis, Planning, Implementation and Control*, 7th ed., Prentice-Hall, Englewood Cliffs, NJ, 1991.
- [17] George, W. R. & Grönroos C. *Developing customer-conscious employee at every level internal marketing*. in Handbook of services marketing. Carole A. Congram and Margaret L. Friedman, eds. AMACOM, New York, 1989.
- [18] Tsai, Ya-Fang, ”Internal marketing and service quality”, *conference proceedings of NDSI , The 2006 Annual Meeting of the Northeast Decision Sciences Institute*, in San Juan, Puerto Rico, 2006.
- [19] Mathieu, J.E., Zajac, D.M. ”A review and meta-analysis of the antecedents, correlates, and consequences of organizational commitment.” *Psychological Bulletin*, 1990, 108(2), 171-194.
- [20] Allen, Natalie J. & Griffeth, Douglas, B. ”Employee commitment to the organization and customer reactions mapping The linkages.” *Human Resource Management Review*, 2001, Vol.11, 209-236.
- [21] Buchanan, B. ”Building organizational commitment: The socialization of managers in work organizations.” *Administrative Science Quarterly*, 1974, Vol.19, 533-546.
- [22] Mowday R.T., Steers R.M, Porter, L.W. ”The measure of organizational commitment.” *Journal Vocational Behavior*, 1979, Vol.14, 224-247.
- [23] Meyer, J.P., and Allen, N.J. ”A three-component conceptualization of organizational commitment.” *Human Resource Management Review*, 1991, Vol.1, 61-89.
- [24] Pfeffer, J. ”Producing sustainable competitive advantage through the effective management of people.” *Academy of Management Executive*, 1995, Vol.9, 55-72.
- [25] Guest, D., & Conway, N. ”Peering into the black hole: The downside of the new employment relations in the UK.” *British Journal of Industrial Relations*, 1999, Vol.37, 367-389.
- [26] Meyer, J.P., & Allen, N.J. ”Testing the “Side-bet theory” of organizational commitment: Some methodological considerations.” *Journal of Applied Psychology*, 1984, Vol. 69, 372-378.
- [27] Kim, Woo Gon, Leong, Jerrold K., Lee, Yong-Ki. ”Effect of service orientation on job satisfaction, organizational commitment, and intention of leaving in a casual dining chain restaurant.” *International Journal of Hospitality Management*, 2005, Vol.24, 171-193.
- [28] Mankjee, Chandra Rekha, Hartzel, Yolanda F., Uys, Iise L. ”The effect of perceived organizational support on organizational commitment of diagnostic imaging radiographers.” *Radiography*, 2006, Vol.12, 118-126.
- [29] Money, A.H. & Foreman, S. ”The measurement of internal marketing: A confirmatory case research.” *Journal of Marketing Management*, 1996, Vol.11, 755-766.
- [30] Michael T. Ewing & Albert Caruana. ”An internal marketing approach to public sector management-The marketing and human resources interface.” *The International Journal of Public Sector Management*, 1999, 17-26.

A STUDY OF TEAM LEARNING ACTIVITIES IN NPD PROJECT TEAMS – THE MEDIATING MECHANISM OF TEAM MEMBERS’ PERCEIVED TRANSACTIVE MEMORY SYSTEMS

Lung-Far Hsieh, Dept. of Business Administration, Chung Yuan Christian University, Taiwan, R.O.C.,
Tel: 886-3-265-5113, E-mail: lungfar@cycu.edu.tw

ABSTRACT

This study applies transactive memory systems (TMS) as the mechanism connecting division of labor with team learning activities. A field study was conducted to verify and refine this framework and to provide potential theoretical explanations for the influences of division of labor at organizational level on its NPD teams’ learning activities. Several propositions derived from this field study are: 1) Higher levels of division of labor at organizational level associate with higher levels of TMS perceived by the team members, 2) Higher levels of responsibility assignment also associate with higher levels of TMS, and 3) An earlier and clearer TMS associates more team learning activities.

Keywords: Team learning activities, Division of labor, Transactive memory systems

INTRODUCTION

Recently, more and more companies have employed cross-functional teams as the primary integrating mechanism in their new product development (NPD) activities (Leenders and Wierenga, 2002). An implicit assumption for the productivity of a cross-functional team is that, when team members with disparate specialties are brought together, there comes a synergistic integration of their different knowledge that will boost the team’s performance. This assumption parallels to Brown and Duguid’s (2001) contention that a firm’s advantage over the marketplace is found in “the greatest benefits to knowledge accrue from coordinating its development across the division of labor”. However, empirical findings from NPD literature don’t seem to consistently support this synergistic integration assumption (Barczak and Wilemon, 2003; Sarin and Mahajan, 2001). Research in small groups’ information processing dynamics (e.g., Stasser, Taylor and Hanna, 1989; Stasser and Stewart, 1992; Stasser et al. 1995; Larson et al., 1994; Stewart et al. 1998) also reveals that the synergistic effects may not come automatically by simply pooling team members with various knowledge.

On the other hand, research in knowledge management and organizational learning has constantly demonstrated a strong association between effective team learning activities and higher team performance (Dougherty, 1992; Moorman and Miner, 1997; Sarin and McDermott, 2003; Katzenbach and Smith, 2003). In this vein, Hage (1999) emphasized the importance of the complexity of the division of labor because it taps the organizational learning, problem-solving, and creativity capacities of the organization. Therefore, it is arguable that the effective team learning activities is highly relevant to the synergistic effect of a cross-functional team. This linkage between division of labor and team learning activities is central to the justification of adopting cross-functional teams in new product development.

Because that most research concerning the influences of division of labor on team activities is largely found in small group dynamics research and in laboratory settings (e.g., Stasser, 1989; Stasser, and

Stewart, 1992; Larson et al., 1994; Stewart and Stasser, 1995; Stasser et al. 2000; Lewis et al., 2005), there is little empirical evidence supporting the linkage between the division of labor in an organization and its cross-functional teams' learning activities. Neither is there a clear theory to explain the underlying mechanism of this linkage. This is lamentable because there is a tradition in the research of the influences of division of labor in product innovation (Damanpour, 1991) and, from the viewpoint of practitioners, the division of labor is one of the initial conditions of cross-functional teams that are mostly amenable to the discretion of managers in an organization.

Research in transactive memory systems (TMS) (Wegner, 1987) demonstrates a promising theory to explain the linkage between division of labor and team learning activities. Studies of team behavior have revealed that effective teams use a well functioning TMS to store and recall more knowledge than any individual (Hollingshead and Brandon, 2003), to use the knowledge of others better (Moreland, Argote and Krishnan, 1998), to match problems with the person most likely to resolve them (Moreland and Levine, 1992), and to reduce cognitive load when others act as external memory stores and allow greater specialization (Hollingshead and Brandon, 2003). Therefore, this study will use transactive memory systems (Wegner, 1987) to explain the linkage between division of labor and team learning activities and provide evidence to support it.

THE FIELD STUDY

The study conducted a field study on team members' perceived TMS by in-depth interviewing members of two NPD project teams from two computer & communication components manufacturing companies in Taiwan. These interviews were conducted at early, middle, and last time periods of NPD stages of each team, respectively. First, the researcher implemented a questionnaire survey on NPD managers of 10 computer & communication components manufacturing companies to investigating levels of division of labor of these companies. The questionnaire consists of three components of division of labor, including specialization, functional differentiation, and professionalism. The two teams were selected purposefully that one team (Team A) came from the company with the highest level of division of labor, and the other (Team B) came from the company with the lowest level of division of labor. Team A consisted of 11 members including seven engineers (three product developing engineers (components engineers), one design engineer, one product structure (layout) engineer, one client service engineer, one senior product developing engineers serving as the leader of the team) and four part-time participants (one purchaser, one manufacture engineer (cost analyzer), one marketing analyzer, and one book-keeper). Team B consisted of 10 members including eight engineers (four product design engineers, three senior product design engineers, one product design manager serving as the leader of the team) and two part-time participants (one purchaser and one adviser from marketing department). The study will focus on investigating the formation of the team members' perceived TMS and its effectiveness for acting an underlying mechanism linking division of labor at organizational level and team learning activities of NPD project teams.

The formation of Transactive memory system

Lewis (2003) demonstrated that the TMS construct have three major components: specialization, credibility, and coordination. Thus, this field study will investigate these TMS components to estimate the level of TMS on the observed teams. On the in-depth interviews at the early NPD stage (conducting right after each team had started for one week), most of team members of both teams reported high awareness of their roles in the NPD teams. But when asked to depict other members' roles, members of

Team A reported teammates' specialized domains more consistent with their self-reported ones than did members of Team B. Two junior engineers of Team B (with less than 2 years in the company) even confused the specialties of their three senior product design engineers. When asked if they had confidence on the specialties of their teammates, members of Team A also reported higher trust on their teammates' specialties than did members of Team B. This difference mainly comes from two senior engineers of Team B who expressed reservations on competence of some of their junior engineers. During some informal conversations when I observed Team B's designing works as a bystander, some junior engineers complained to me about being devaluated by their senior teammates.

On the in-depth interviews at the middle NPD stage conducting right after each team had reached the middle point of its outline schedule (the outline schedule of Team A and Team B is eight months and six months, respectively), all members of both teams reported high awareness of their roles in the teams. But some members of Team B still revealed certain confusions and some still expressed reservations of their teammates' specialized domains, although much less than they did at the early stage of the NPD program.

By reaching its middle of the fourth month from the starting of NPD program, Team B's manager redesigned the NPD program and prolonged its schedule to ten months because of the depressed progression of the NPD tasks. Beside, he also explicitly defined the four junior engineers' responsible specialties that set the responsibilities for knowledge of certain domains to each junior engineer. Adapting to this change in Team B, this study conducted an additional in-depth interviews with members of Team B right after five months from its beginning. This time, all members of Team B reported high awareness of their roles in the team and predicted teammates' specialized domains that consisted with their self-reported. Moreover, they also expressed high level of confidence on teammates specialties, although a little less than that of Team A.

On the in-depth interviews at the last NPD stage (conducting right after each team had reached the beginning of the last month of its schedule), all members of both teams reported high awareness of their roles in the teams, depicted consistent specialized domains of teammates, and expressed high level of confidence on teammates specialties. Thus, in this stage, the teams' TMSs had a stable condition.

The linkage between division of labor and TMS

In this field study, Team A came from the company with the highest level of division of labor among ten investigated companies. It also exhibited much higher levels of TMS than did Team B, especially in early stage of the NPD program. Just after one week from beginning, members of Team A already worked smoothly with one another. This provides some temporal evidence for the positive association between division of labor in an organization and the level of its team's TMS. This also confirms Lewis's (2004) argument that teams with initially distributed expertise and familiar members are more likely to develop a TMS and "TMS emerge during a project-planning phase as a function of a team's initial conditions". Hence, the study puts the proposition as following:

P1: a NPD team with higher level of division of labor is more likely to form its TMS.

There is some unanticipated finding that, after Team B's manager explicitly assigned responsibilities on team members for certain knowledge domains, Team B seemed able to develop its TMS effectively. This resonates to a recent study that emphasizes the importance of allocating responsibilities for knowing within a team (Jackson and Klobas, 2008). But, this administrative coordination on members'

expertise seems to conflict with Faraj and Xiao's (2006) observation that expertise coordination practices are highly emergent and cannot necessarily be pre-specified. And this study also found that although Team B's level of informal discussion (one of learning activities) significantly increased, its level of reflection (another one of learning activities) was hardly heightened. Nevertheless, this study indeed observed an association between responsibility assignment among team members and the team's TMS. Thus, I propose the following.

P2: a NPD team with higher level of responsibility assignment is more likely to form its TMS.

Team learning activities

In organizational learning literature, two major processes of learning are learning-by-absorption and learning-by-reflection (Scarborough et al. 2004). Accordingly, team learning should not be confined to knowledge transfer among team members, it also assumes combining all members' knowledge and experiences to cope with new situations. For cross-functional teams, combining (or, integrating) team members' separate knowledge entails negotiating and reflecting along their underlying backgrounds such as assumptions, routines, mental models, etc. In this vein, Edmondson (2002) identified team learning as the iterative process of reflective discussion and action. Only through an iterative cycle of reflective discussion and action, can a team truly integrate specialist capabilities of all team members.

This study defines team learning as an iterative process in which a team's members collectively gather knowledge from themselves, their internal colleagues (intra-organizational) and external sources (inter-organizational); engage the process of discussion and reflection to make decisions to cope with the assigned common tasks; and gain feedbacks and additional knowledge to embark on next round of knowledge gathering, discussion and reflection. Following this definition, team learning activities are those activities (taken by members of a team) that conducive to team learning, including knowledge gathering, discussion, and reflection. Thus the field study focuses on these activities of team members.

From beginning, members of Team A demonstrated much higher level of knowledge gathering than did members of Team B. Usually, some members of Team A would provided their newly gained information in the routine meetings. Sometimes, they even brought in certain colleagues in other departments to provide first hand knowledge, especially on some bottlenecks they confronted in later stages of the NPD program.

There also were many informal discussions engaged among variant compositions of team members. The conversations often started by two team members discussing their design problems, then someone was asked to join the discussion, still then may be joined by more teammates. These phenomena continued to occur from the early stage to last stage of the NPD program in Team A. But, in Team B, the informal discussions of designing works are much less in the early and middle stages of the NPD program. Only during the later stage did members of Team B frequently engaged this kind of discussion, although still less frequent than those occurred in Team A. In contrast, the formal discussions (estimated as time spent on routine meetings) in Team B were much longer than those in Team A. Interestingly, in Team B, the time of routine meetings seemed to decrease during the later stage of NPD program as its informal discussions increased.

In interviews, members of Team A expressed more eagerness for negotiating potential conflicts relating to their responsible design areas with other teammates. Basically, they thought that both informal discussions and routine meetings were used to negotiate these interdependences and search for some

alternatives. While in Team B, the routine meetings were deemed to allocate design responsibilities and inspect if tasks were done in schedule. But when asked if they thought that reflection is the important mean for integrating team members' different perspectives, members of both teams all highly confirmed with it. When asked if their meetings' conclusions were decisive, members of team B expressed negative answers (by average) in interviews at early stage and first middle stage (three months after beginning) and expressed positive answers (by average) at second middle stage (five months after beginning) and last stage of the NPD program. While members of Team A expressed positive answers at all stages of NPD program.

For more fine-grained analyses, it is worth to mention that there is a subtle difference between reflection and discussion. According to social loafing research (Karau and Williams, 1993) there is a tendency for people to expend less effort when working collectively than when working individually. In most of cross-functional teams, there is also a tacit behavioral norm that demands team members to engage in team discussions. As a result, team members may participate discussions but never touch critical issues. In this field study, members of Team B spent much time in routine meetings, but these meetings often resulted in less decisive conclusions. In other words, they may engage the form of discussions but provide little reflection.

The linkage between TMS and team learning activities

In this field study, members of Team A exhibited much higher level of knowledge gathering than did members of Team B. This is consistent with Ren et al.'s (2006) finding that TM decreases group response time by facilitating knowledge retrieval processes and improves decision quality by informing task coordination and evaluation. Also, the many informal discussions within parts of members further speeded the efficiency of knowledge transfer among members with different specialties. There are many studies (Stasser et al., 1995; Littlepage et al., 1997; Moreland and Myaskovsky, 2000) that provide consistent evidence for the effectiveness of TMS on enhancing team leaning activities.

Members of Team A also expressed an eagerness for negotiating potential conflicts relating to their responsible design areas with other teammates. Brown and Duguid (1998) argued that tacit knowledge could be communicated or demonstrated only in direct interaction, and corrected in mutual adjustment between deliverer and recipient, therefore to combine different experts' experiences needs the process of iterative reflection and action. These activities will make team members to suspend assumptions and enter into a genuine "thinking together" (Senge, 1990), which is critical to a team's creative capabilities. Thus, this study proposed that:

P3: an earlier and clearer TMS will result in more team learning activities.

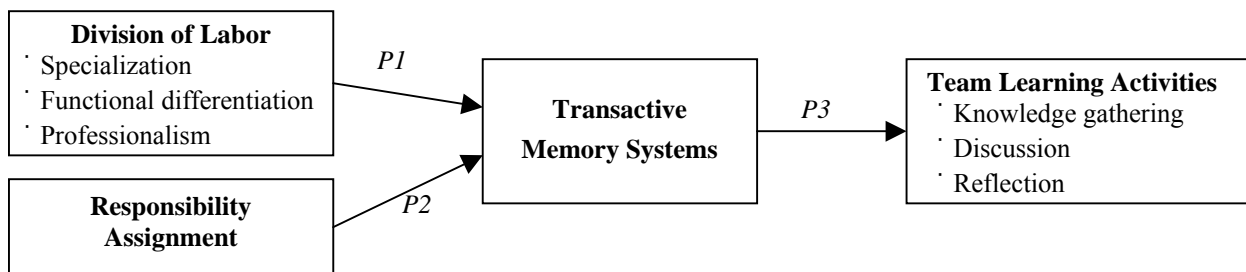


Figure 1. A Summary of Research Propositions

CONCLUSIONS

In a multi-functional team consisting of experts with different specialties, team members often bring with significant differences in perspectives and interests (Carlile, 2002). It is necessary to provide some mechanisms for facilitating the transformation and combination of different knowledge domains within these teams. As Faraj and Sproull (2000) put it that to manage members' skill and knowledge interdependencies effectively entails knowing where expertise is located, knowing where expertise is needed, and bringing needed expertise to bear. The theory of transactive memory system (TMS) seems to provide a plausible explanation to fulfill this requirement. This study provides some field materials to temporarily support and refine the theorizing that TMS acts as a mediating mechanism to link division of labor at organizational level with team learning activities of the NPD team. Further empirical studies incorporating more samples and applying more refined measures should provide more decisive findings on the research of TMSs.

REFERENCES

- [1] Leenders, M.A., Wierenga, B., "The effectiveness of different mechanisms for integrating marketing and R&D," *Journal of Product Innovation Management*, 2002, 19(4), 305-317.
- [2] Brown, J.S., Duguid, P., "Knowledge and organization: A social-practice perspective," *Organization Science*, 2001, 12(2), 198-213.
- [3] Braczak, G., Wilemon, D., "Team member experiences in new product development: Views from the trenches," *R&D Management*, 2003, 33(5), 463-479.
- [4] Sarin, S., Mahajan, V., "The effect of reward structures on the performance of cross-functional product development teams," *Journal of marketing*, 2001, 65(2), 35-53.
- [5] Lewis, K., "Measuring transactive memory systems in the field: Scale development and validation," *Journal of Applied Psychology*, 2003, 88(4), 587-604.
- [6] Jackson, P., Klobas, J., "Transactive memory systems in organizations: Implications for knowledge directories," *Decision Support Systems*, 2008, 44(1), 409-424.
- [7] Edmondson, A., "The local and variegated nature of learning in organizations: A group-level perspective," *Organization Science*, 2002, 13(2), 128-146.
- [8] Hage, J.T., "Organizational innovation and organizational change," *Annual Review of Sociology*, 1999, 25, 597-622.
- [9] Reagans, R., Argote, L., Brooks, D., "Individual experience and experience working together," *Management Science*, 2005, 51(6), 869-881.

A complete form of references is available on request.

A Comparison of Affiliative Organizational Citizenship Behavior and Challenging Organizational Citizenship Behavior

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

ABSTRACT

Based on the classification of extra-role behaviors from Van Dyne, Cummings, and McLean Parks (1995), this study proposes that organizational citizenship behaviors (OCBs) should be distinguished into two types, affiliative OCBs and challenging OCBs, as each type has its specific characteristics. In this paper, the differences between affiliative OCBs and challenging OCBs are examined and several contentions are proffered. First, in regard to the behavioral motivations, affiliative OCBs may be more associated with instrumentality beliefs, while challenging OCBs may be more associated with moral beliefs. Second, affiliative OCBs and challenging OCBs may have different sets of antecedents. Finally, affiliative OCBs are more likely to result in positive consequences for individuals than are challenging OCBs.

INTRODUCTION

Organizational employees are expected to not only follow written rules and job descriptions, but also spontaneously perform behaviors which go beyond delineated role expectations and also benefit the organization (Van Dyne, Graham, & Dienesch, 1994). Organ and his colleagues referred to these spontaneous behaviors as organizational citizenship behaviors (OCBs) (Organ & Bateman, 1983; Smith, Organ, & Near, 1983). The OCB construct has received a lot of research attention. Until now, scholars have identified numerous forms of citizenship behaviors (Podsakoff, MacKenzie, Paine, & Bachrach, 2000), and empirically verified its positive impacts on organization effectiveness and individual performance evaluation (Avila, Fern, & Mann, 1988; MacKenzie, Podsakoff, & Fetter, 1993; Podsakoff, & MacKenzie, 1994; Van Scotter & Motowidlo, 1996; Walz & Niehoff, 1996).

While the OCB research stream has accumulated prolific results, questions regarding its notion and content have arose in recent years. For example, researchers have mentioned that the OCB classification is not complete, and that some forms of behaviors are neglected, especially those behaviors challenging the present state of organizations (Morrison & Phelps, 1999). Van Dyne, Cummings, and McLean Parks (1995) suggested that extra-role behaviors could be divided into affiliative behaviors and challenging behaviors. The former behaviors are other-oriented and promote cooperation, while the latter behaviors are changed-oriented and aim at improvement. Generally speaking, most forms of OCBs examined within the existing research are affiliative behaviors. Since challenging behaviors also are a part of employees' spontaneous behaviors and contribute to the effectiveness of organizations, this study contends that challenging behaviors should be included in the domain of OCB.

Further, most OCB studies have tended to adopt a supervisory perspective and focused on the benefits to work units or organizations; few studies have tried to understand its potential implications for employees (Bolino & Turnley, 2005). Sometimes, behaviors that comply with supervisor opinions may be appreciated even though these behaviors are quite impression-enhancing or self-serving, while behaviors which are not in accordance with supervisor opinions may receive negative feedback even though these behaviors are quite altruistic. In the long term, managers and organizations not only need employees who are willing

to comply with their policies, but also employees who are willing to provide constructive critiques.

The objective of this study is to conduct a theoretical comparison on affiliative OCBs and challenging OCBs in terms of four aspects: instrumentality beliefs, moral beliefs, antecedents, and consequences. Since previous research on challenging OCBs is still quite rare, such a theoretical comparison may provide insights for the development of this research stream, and provide a knowledge base for managers to adopt effective managerial skills or suitable motivational methods to encourage such kinds of behavior.

THE FORMS OF AFFILIATIVE OCB AND CHALLENGING OCB

According to a literature review compiled by Podsakoff et al. (2000), almost 30 potentially different forms of citizenship behavior have been identified. Although there is a great deal of conceptual overlap, these forms of citizenship behavior can be organized into seven common themes or dimensions: (1) helping behavior (altruism), (2) sportsmanship, (3) organizational loyalty, (4) organizational compliance, (5) individual initiative, (6) civic virtue, and (7) self development.

Van Dyne et al. (1995) proposed that extra-role behaviors could be categorized into two types: affiliative behaviors, which tend to solidify or preserve the interpersonal relationships in organizations, and challenging behaviors, which propose to change the present states and bring about improvements. Extant classifications of OCBs emphasize the affiliative type behaviors. For example, among the seven dimensions in Podsakoff et al. (2000), helping behavior, sportsmanship, organizational loyalty, and organizational compliance are typical affiliative behaviors, while only a small part of civic virtue behavior (such as expressing opinions) can be seen as a type of challenging behavior.

Although many challenging extra-role behaviors have not yet been formally included in the research domain of OCB, these behaviors do conform to the tradition definition of OCBs, that is, these behaviors are discretionary, go beyond existing role expectations, and are intended to benefit the organization (Van Dyne et al., 1995). Therefore, this study refers to these behaviors as challenging OCBs. In the literature, several constructs related to the challenging OCB construct exist such as taking charge (Morrison & Phelps, 1999), issue selling (Dutton, Ashford, O’Neill, & Lawrence, 2001), voice (Rusbult, Farrell, Rogers, and Mainous, 1988), advocacy participation (Van Dyne, Graham, & Dienesch, 1994), and principled organizational dissent (Graham, 1986). A description of these constructs is listed in Table 1.

Table 1 Descriptions of Different Forms of Affiliative OCB and Challenging OCB

AFFILIATIVE OCB	
Behavioral Form	Description
Helping behavior	Voluntarily helping others with or preventing the occurrence of work-related problems (Podsakoff et al., 2000).
Sportsmanship	A willingness to tolerate the inevitable inconveniences and impositions of work without complaining (Organ, 1990).
Organizational Loyalty	Endorsing, supporting, and defending organizational objectives (Borman & Motowidlo, 1993).
Organizational Compliance	Following organizational rules and procedures (Borman &

CHALLENGING OCB

Behavioral Form	Description
Voice	Actively trying to improve conditions; actively searching for and coming up with new ways of doing things and advocating changes to make things better (Zhou & George, 2001).
Issue selling	Affecting others' attention to and understanding of the events, developments, and trends that have implications for organizational performance (Dutton, Ashford, O'Neill, & Lawrence, 2001).
Taking charge	Employees' voluntary and constructive efforts to effect organizationally functional change with respect to how work is executed within the contexts of their jobs, work units, or organizations (Morrison & Phelps, 1999).
Advocacy participation	Maintaining high standards, challenging others, and making suggestions for change (Van Dyne, Graham, & Dienesch, 1994).
Principle organizational dissent	The effort by individuals in the work-place to protest and/or to change the organizational status quo because of their conscientious objection to current policy or practice (Graham, 1986).

**INSTRUMENTALITY BELIEF AND MORAL BELIEF
ON AFFILIATIVE OCB AND CHALLENGING OCB**

In the literature on affiliative OCBs, there is an implicit assumption that the current policies advocated by managers or the current practices conducted in organizations are adequate, correct, or function properly. Hence, employees who engage in behaviors complying with status quo will benefit their work units, and will help lead to overall organizational effectiveness. However, managers and organizations are often faced with problematic events which call for sweeping changes in order to adapt to complex environments. Thus, in some circumstances, managers and organizations may find it beneficial to retain employees who are willing to express new ideas or constructive criticisms.

From the employees' standpoint, such OCB displays may enhance their personal performance evaluations or influence within the company, but these types of behaviors may also result in personal costs or disadvantages. For example, affiliative OCB may increase an employee's role overload (Bolino & Turnley, 2005), while challenging OCB may offend the authority within the organization (Graham, 1986). Since employees often perceive OCBs as a kind of trade-off or paradox, whether to engage in OCBs may be in some way associated with their instrumentality beliefs and moral beliefs regarding these types of behavior. Both of these beliefs are discussed in the following section.

Instrumentality Beliefs

Recent research suggests that employees believe that managers fairly reward their OCBs (Haworth & Levy, 2001). Although the original OCB definition treats these behaviors as purely discretionary and not formally rewarded, empirical results have shown that OCBs are positively related to managerial performance evaluations (Podsakoff, MacKenzie, & Hui, 1993). Affiliative OCBs are usually in line with managerial expectations, so these behaviors are likely to be rewarded and helpful in terms of enhancing one's individual image. Further, Bolino (1999) proposed that employees may be motivated to engage in OCBs for impression-management

reasons. However, it is quite uncertain whether challenging OCBs will be favored by managers, especially in terms of the behavior that goes contrary to managerial points of view. Employee contributions regarding questionable policies or inappropriate procedures may result in rewards from conscientious managers, but may lead to offense with more narrow-minded managers. Therefore, challenging OCBs may not be directly linked to favorable consequences or positive feedback. While the performances of affiliative OCBs and challenging OCBs both may lead to real disadvantages (such as increased workload or sacrifices in terms of individual task performance), overall, the negative consequences of challenging OCBs may be greater than those of affiliative OCBs. Thus, the relationship between instrumentality beliefs and affiliative OCBs may be stronger than the relationship between instrumentality beliefs and challenging OCBs.

Moral Beliefs

The metaphor of “citizenship” implies that employees may engage in OCBs because of a sense of obligation or altruistic motivation; hence previous researchers have described OCB performers as organizational “good soldiers” (Bateman & Organ, 1983). Such a theoretical point suggests that the motivations of OCB may lie in employees’ moral reasoning or moral beliefs about what makes the behavior good (Ryan, 2001). Kohlberg (1969), a developmental psychologist, proposed a stage theory describing the development of moral reasoning. Low stages (pre-conventional level) of moral reasoning are characterized by instrumental egoism and the morality of obedience. Middle stages (conventional level) emphasize compliance to law and duty so as to insure social order. High stages (post-conventional level) emphasize the underlying principles that serve as foundations for law (Ryan, 2001), thus the type of moral reasoning in these stages can also be labeled as “principled moral reasoning”.

Ryan’s (2001) study found that principled moral reasoning is positively related to the OCB dimensions of interpersonal helping and sportsmanship, but not positively related to the dimension of civic virtue. Since the other OCB dimensions have not yet been examined, the contention that moral reasoning serves as a determinant of OCBs has not been fully verified. It is likely that affiliative OCBs and challenging OCBs are based on varied stages of moral reasoning or moral beliefs. For example, “complying with organizational policies” may be more related to conventional moral reasoning because such kinds of behavior aim at maintaining the organizational order or rule; while “challenging the morality of authority figures” may be more related to post-conventional moral beliefs because such kinds of behavior may be related to an intent pertaining to the universal justice or ethics transcending the organizational criterion. However, because both types of OCB have a variety of forms, it is likely that different forms of behavior are characterized by different types of moral reasoning, and therefore can not be treated homogeneously. Accordingly, the current study proposes that the relationship between principled moral reasoning and challenging OCBs may be stronger than the relationship between principled moral reasoning and affiliative OCBs.

ANTECEDENTS OF AFFILIATIVE OCB AND CHALLENGING OCB

According to Podsakoff et al. (2000), empirical research has identified four major categories of OCB antecedents: employee characteristics (e.g., including satisfaction, organizational commitment, agreeableness, etc.), task characteristics (e.g., including task feedback, task routinization, etc.), organizational characteristics (e.g., organizational formalization, perceived organizational support, cohesive group, etc.), and leadership behaviors (e.g., transformational leadership, leader-member exchange, etc.). While research has identified numerous OCB antecedents, most studies focus on exploring the antecedents of affiliative OCB, whether these antecedents can effectively predict challenging OCB is unknown. For instance, job satisfaction

may be an important antecedent requiring further investigation. While research has found a positive relationship between job satisfaction and some forms of affiliative OCBs (such as altruism, conscientiousness, and sportsmanship), the relationship between job satisfaction and challenging OCBs is indeterminate. Employees with high job satisfaction may not want to change the status quo, thus job dissatisfaction could prove to be a driving force for employees who opt to challenge their organizations (Rusbult, Farrell, Rogers, & Mainous, 1988). However, extremely high level of job dissatisfaction may make employees choose to exit or keep silent rather than take constructive action (Withey & Cooper, 1989). Since there is no consensus on the relationship between job satisfaction/dissatisfaction and challenging behaviors, some scholars have called for an investigation of the contingent variables in such a relationship (Janssen, Vries, & Cozijnsen, 1998).

Further, due to their differing characteristics, affiliative OCBs and challenging OCBs may be influenced by different sets of antecedents. Van Dyne, Cummings, and McLean Parks (1995) suggested that some factors rarely noted in the literature pertaining to affiliative OCBs, may be related to challenging behaviors, such as: propensity to take risks, internal locus of control, need for power, moral development, and career involvement. Thus, in order to advance our understanding of challenging OCBs, an exploration of the antecedents of challenging OCBs may follow one of two approaches. First, researchers may examine the same set of antecedents in order to understand whether these antecedents have similar functions in terms of affiliative OCBs and challenging OCBs. Second, researchers may attempt to identify and examine a new set of antecedents which have not yet been identified in the previous affiliative OCB literature.

CONSEQUENCES OF AFFILIATIVE OCB AND CHALLENGING OCB

The literature usually takes a positive view regarding the consequences of affiliative OCBs. In regards to organizational-level outcomes, helping behavior and sportsmanship have been found to be positively related to organizational performance (Podsakoff, Ahearne, & MacKenzie, 1997; Walz & Niehoff, 1996). As for individual-level outcomes, most forms of affiliative OCBs have been found to positively influence subsequent managerial performance evaluations (Podsakoff et al., 2000). While the majority of past studies have emphasized the positive features of OCBs, Bolino, Turney, and Niehoff (2004) listed potential negative aspects of such behaviors. For example, they suggested that OCBs may be indicative of poor management or understaffing, so that employees are frequently called upon to work long hours or to continually help one another. In addition, employees may feel pressured to continually increase their acts of citizenship in order to be seen as a good organizational citizen, which may leave them feeling stressed and overloaded (Bolino & Turnley, 2005). Thus, it is also possible that affiliative OCBs may result in unfavorable consequences to the organizations or the individuals involved.

Empirical research has not yet focused any attention on the consequences of challenging OCB. According to Van Dyne et al.'s (1995) perspective, challenging behaviors may lead to positive organizational consequences. For example, the behavior of issue selling may enhance the innovation and the behavior of principled dissent may prevent immoral or inappropriate organizational practices. However, the individual-level consequences of challenging OCBs are quite uncertain. When an employee expresses dissent, he/she may do so according to a sense of doing the "right" thing (a positive internal feeling), but may in turn face hostility from managers or peers. It is also likely that an employee will feel a sense of contribution following the effective execution of his/her idea, or will positively receive compliments or a promotion from his/her manager. Since the relationship between challenging OCBs and individual or organizational outcomes can be very intricate, identifying any moderating variables could certainly assist in

discovering their true nature. For example, Dutton, Ashford, O'Neill, and Lawrence (2001) revealed that employees' contextual knowledge is critical to the success of issue-selling behaviors.

CONCLUSION

According to the theoretical classification proposed by Van Dyne et al. (1995), this study proposes to distinguish OCBs into two types: affiliative OCBs and challenging OCBs. While OCB research has produced fruitful results, most of the studies have focused on affiliative OCBs. Therefore, the results may not apply or may not be generalizable to challenging OCBs. In order to enhance our understanding of challenging OCBs, this study conducted a preliminary theoretical comparison of these two types of OCBs. Based on a literature review, I suggest that affiliative OCBs and challenging OCBs have a great deal of differences related to aspects of instrumentality beliefs, moral beliefs, antecedents, and consequences. First, affiliative OCBs may be more associated with instrumentality beliefs, as this kind of behavior is more likely to result in positive individual outcomes; however, challenging OCBs may be more associated with moral beliefs because this kind of behavior is less likely to lead to strictly positive feedback. Second, affiliative OCBs and challenging OCBs may be influenced by different sets of antecedents. Third, it is easier for an employee to predict potential positive outcomes when he/she exhibits affiliative OCBs than when he/she exhibits challenging OCBs.

Given the importance of challenging OCBs to the effectiveness of organizations, researchers should pay more attention to the antecedents and consequences of this kind of behavior. Such an examination could provide a knowledge base for managers to adopt adequate managerial skills or design suitable systems to encourage employees to engage in these types of behaviors. Further, the extant literature reveals that the causal relationships between challenging OCBs and other variables (including antecedents and consequences) are very intricate; thus, an important task for future research is to identify relevant contingent factors which may help clarify the above relationships.

REFERENCES

References available upon request from the author.

THE INFLUENCE OF INSTITUTIONS AND TRANSACTION COST ON CHANNEL STRUCTURE: THE CASE OF ASUS IN MAINLAND CHINA

Chi-Hsing Tseng, National Pingtung Institute of Commerce,
No. 51, Min Sheng E. Rd., Pingtung 900, Taiwan, R.O.C.
886-8-7238700 Ext. 6215
e-mail: tseng@npic.edu.tw

Hsin-Chih Kuo, I-Shou University,
No. 1, Sec. 1, Syuecheng Rd., Dashu Township, Kaohsiung County 840,
886-7-6577711 ext.6461
e-mail: simon@isu.edu.tw

ABSTRACT

This study integrates institutional theory and transaction cost theory to explore the marketing Channel of Taiwanese Notebook PC firms in Mainland China. The purpose of this study is to find out the evolution of channel structure under both pressures of institution and efficiency. After surveying the case of Asus, this research finds that Taiwanese firms' marketing channel tend to be isomorphism alliance as the local institutional environmental pressure is strong. As the institutional environmental pressure becomes weaker and the transaction cost is relatively high, the Taiwanese firms' marketing channel tends to be internalization. As the institutional environmental pressure becomes weaker but the transaction cost is relatively low, the Taiwanese firms' marketing channel tends to be efficiency alliance.

Keywords: Institutional theory, Transaction cost, Channel structure, ASUS, Mainland China

INTRODUCTION

Since the 1980s, the amount of FDI by Taiwanese corporations has grown enormously. Among these outward investment countries, China represents the most popular host countries. As the Chinese domestic market has been opened, the investing activities of Taiwanese firms have been transferred gradually from building production base to exploring local market. How to manage business in China has been the concern focus of practitioners. However, Chinese market possesses the environment uncertain traits of emerging market. The transaction cost and relevant resources in China were then different from the West. Recently, the number of scholars applying institutional theory to analyze the strategies of emerging market was increasing gradually.

Different schools have different interpretation about the formation, evolution, operation, and outcomes of organizational alliances and networks. From the view of theory development, there are two streams of thought: economics-based views of alliances and the corporate strategy perspective. Most previous studies based on transaction-cost theory have tended to understate the significance of contextual factors in the choice of entry mode. On the other hand, many of these contributions from corporate strategy perspective lack a cohesive theoretical perspective. The institutional perspective offers the hope of potential integration of the economic and strategic perspectives because of the multiplicity and complexity of alliances and networks. After that, there is a growing number of studies began to apply institutional theory to analyze international business management. Besides, marketing channel can be viewed as relationship network, therefore the study treated marketing channel as an alliance network.

THEORETICAL BACKGROUND

Transaction Costs and Channel Structure

In a general sense, transaction cost theory views governance in terms of designing particular mechanisms for supporting economic transactions. The original framework, as developed by Williamson (1975), views the governance decision as fundamentally a choice between a market and a hierarchy. Williamson (1991) argued that whether transactions are conducted within the market, internalized within the firm or conducted alternatively by an intermediate form of flexible supply, contracting is determined by the governance mechanism that minimizes transaction costs, and ultimately maximizes efficiency. The greater the uncertainty, the higher the transaction specific investments involved and the stronger the need for vertical integration (Williamson, 1985). Transaction cost theory is useful to explain the organizational form of exchange between partners in an export marketing channel (Heide, 1994; Tesfom, Lutz, Ghauri, 2004). The higher the transaction cost, the more likely the firm will adopt internalized marketing channel structure.

Institutions and Channel Structure

Recent advances in organization theory suggest that organizations strive for both economic fitness, which emphasizes the competition for scarce resources and underscores the importance of the task environment, and social fitness, which stresses the pursuit of legitimacy in the eyes of important societal stakeholders and accentuates the significance of the institutional environment (DiMaggio & Powell, 1983; Oliver, 1991). Grewal and Dharwadkar (2002) incorporated the institutional environment into current marketing channels research. The institutional environment perspective relies on the primacy of regulatory institutions, normative institutions, and cognitive institutions in influencing the legitimacy of channel members in the larger societal context. Legitimacy is a generalized perception or assumption that the actions of an entity (channel member) are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions (Suchman, 1995). Unlike the task environment approach, with its emphasis on uncertainty and dependence reduction, the institutional approach focuses on the necessity of organizational legitimacy. This concern with social fitness leads to the development of processes that result in the formation of institutions and the emergence of corresponding institutional mechanisms that influence the internal polity and economy of marketing channels (Grewal & Dharwadkar, 2002). Grewal and Dharwadkar (2002) extend the political economy framework by developing three institutional processes (regulating, validating, and habitualizing) that lead to the formation of types of institutions (regulatory, normative, and cognitive, respectively). They suggest that the institutional processes result in institutional pressures, which in turn influence channel structures and processes.

The Integrated Model

Transaction cost theory and institutional theory, two established perspectives within organization studies, offer seemingly contradictory interpretations of organizational phenomena. Roberts and Greenwood (1997) concluded that economic approaches to the study of organization, transaction cost analysis included, generally focus on efficiency, whereas institutional theorists place particular emphasis on legitimation processes and the tendency for institutionalized organizational structures and procedures to be taken for granted, regardless of their efficiency implications. In their research, transaction cost and institutional perspectives are not necessarily in conflict, but are complementary elements of a constrained-efficiency framework. Martinez and Dacin (1999) provided a synthesis of transaction cost economics and institutional theory. They presented an illustrative model to explore conditions under which one or both theories may be most appropriate in explaining decision behavior, focusing on two

important situational factors—the degree of ambiguity surrounding transaction cost analyses and the organization’s temporal survival orientation.

THE INSTITUTIONAL ENVIRONMENT OF MAINLAND CHINA

There are three pillars of institutions. However, this research found that there’s little influence of normative institution on notebook PC industry. We will describe the contents of regulative and cultural-cognitive institutions as follow.

Regulative Institution

There are three major regulations that bring pressures to firms’ operations:

- **Domestic Sales Rights:** In the beginning stage of China reform, government set up the constraints of domestic sales percentage to protect domestic industries. The foreign investors need to report the sales percentages of foreign market and China local markets before setting up the enterprise. The report need to submitting to county government of the enterprise location and its upper government levels. The procedures to get domestic sales rights are very complex and also take long time.
- **Government Procurement Law:** China always aims to foster their domestic companies and industries. In year 2003, they have published the “Government Procurement Law” to regulate government procurements should always consider buying domestic goods, engineering jobs and services in higher priority so that they can foster local companies. As to Taiwanese companies, besides the good quality and service of products Taiwanese companies also need to establish good relationship with China governments and cooperate with China companies to get the bids of China governments.
- **Value Added Tax:** China government has regulated the scopes of Value Added Tax (VAT) which include consign the goods to other parties for selling, as well as sell the consigned goods. If Taiwanese companies ask agents to sell merchandises in China, it needs to be levied 17% VAT according to the laws. Even though VAT can transfer to agents then agents transfer to consumers, some China agents do not obey the laws and not pay the taxes. Taiwanese companies need to oversee the agents to pay the VAT to avoid potential lawsuits which will increase the extra costs of Taiwanese companies.

Cultural-Cognitive Institution

Finally, there are two major pressures belong to cultural-cognitive institution:

- **Guanxi:** For doing business in China, it is very important to have personal relationship or marital relatives’ relationship with relevant organizations.
- **Solve conflicts by informal mechanism:** Most business strives in China are reconciled by related parties’ private discussions and not through courts of law. The law structure in China normally allow associations, family organizations and other non-government organizations to solve the contract strives and decide when can terminate cooperation agreements

LINKING THEORY AND PRACTICE: THE CASE OF ASUS

In 1997, after CEO Zon-Tan Shih of Asus surveyed the Chinese market; Asus set up the direction to enter the Chinese market rapidly. Shi-Ming Hsu was responsible for Asus marketing at that time and become the G.M. of Asus China. Hsu brought US dollar two million, and set up an office in Shanghai.

He succeeded to establish the brand renown of Asus mother boards by traveling around all areas of China and giving promotion and speeches in campuses. In 1998, Asus decided to have its own brand in China's markets which includes notebook PC, mobile phone etc. Asus also realized that they must establish their own marketing channels to building up their brand renown. Hsu had been traveling around all big cities in China to set up the marketing channels and built up relationships with channels members. In year 2003 Asus set up their sales network and formally introduced Asus brand notebook PC on to markets.

Since Asus in the Chinese market started with mother boards, it is natural that they sold notebook PC also through these original mother board's agents who were the leading companies in the DIY market. However, Asus China G.M. Yu-Ja Hsu found out that this kind of channel only worked for notebook PC at the beginning stage. After the initial stage, since the product variances were too large and target customers were also different the original sales channel can not achieve the sales target of notebook PC. The competitive advantages of these mother boards agents only exist in their located areas, but can not spread to second, third and fourth level market regions. In summary, the market penetration power of these channels is weak. Moreover, policy constraints in China make it difficult for the foreign computer companies to enter corporate markets especially through the mother board's channels.

To make up the problems happened in regional and corporate markets; in year 2002 Hsu Yu-Ja launched a code name "sky network" plan to re-structure the channels strategy. First, it adopted the principle of not competing with downstream agents. Asus established a "notebook PC selling alliance" which consists of 170 agents who are located in tier 1 or 2 cities. These 170 agents were chosen from the 1,500 companies who cooperate with Asus in long term basis. Moreover, Asus set up 300 Asus notebook PC specialized stores which targeted family consumers and SMB markets and were managed by Asustek Computer Inc. China.

In 2003, Asus started to adopt the nationwide agent approach which they have been using in Taiwan for quite a while. They select Digital China as their nationwide agent and lead their more than six thousand stores network to sell Asus products in whole China. Since then, the downstream agents of Asustek Computer Inc. China operate in specialized stores and the downstream agents of Digital China operate in alliance stores. By the end of 2003, Asus and Digital China built up a sales network alliance and all the agents of both sides from then on could get goods supplied and full support from either party.

At the beginning of 2004, Asus set up "Products and Channels Center Div." whose only work was to assist Digital China sell Asus products well. In Beijing, Asus had a team specialized to coordinate with Digital China. In the seven regions of China, Asus assigned at least one specialist in each region to assist the salesman of Digital China for channel establishment and increasing sales volume.

To expand its market share in 2006, Asus looked for a new alliance. At the end of 2006, Highly Information Industry Co., Ltd. has become the 2nd nationwide agent of Asus. Highly Information Industry Co., Ltd. major sell Asus's AMD notebook PC. Asus cooperates with Highly Information Industry Co., Ltd. because of their professional selling ability in notebook PC. Asus cooperates with Digital China because of their synergy selling ability. In other words, the competitive advantages of Digital China are their profound resources and the competitive advantages of Highly Information Industry Co., Ltd. are their professional knowledge.

THE DRIVING FORCES OF CHANNEL EVOLUTION

Institutional Pressures

Asus did not have domestic sales rights so that they must cooperate with Chinese domestic channel companies. After getting the domestic sales rights, it could execute its strategies according to normal corporate practices. At the second stage, the institutional pressures on these two companies became not so strong. Asus in 2003 entered the corporate market which was still under the constraints of regulative institution so that Asus still needed to cooperate with local agents. For this reason, we can assume that Asus' sales network still faces stronger institutional pressures. However, in the fourth and fifth stages the regulative institution toward Asus became weak. After spending long time and efforts in China, Asus has gradually established good relationship with local channel companies (e.g. Digital China) and set up a channel division. Up to this time, Asus was thoroughly nationwide agents and regional distributors to build up relationship with local downstream dealers and transfer the culture and institutional pressures to nationwide agencies and regional distributors. For the above reasons, we can assume the sales networks of Asus face weaker institutional pressures in the fourth and fifth stages.

Transaction Costs

At the time Asus first entered the China market to sell notebook PC, it was not familiar with the local environment. Plus in China very often people can decide the explanations of laws; the transaction costs are high. To lower the transactions costs of risks and searching, Asus chose to cooperate with past mother board distributors. Before selling notebook PC, Asus had cooperation experience with local distributors in DIY market, so their transactions cost was lower. In the first stage of sales network cooperation processes, Asus found out the mother board channels did not work for selling notebook PC. For these reasons, Asus stopped the original cooperation relationship. To re-gain the domestic sales rights, it needed to search for new cooperation counterparts and negotiate with them. For these reasons, transactions costs in second stage were relatively high. In the third stage, the difficult subject Asus faced was the geographies variances causing the very high management cost and Asus also searching for the most optimum sales network pattern. In third stage, transactions costs were still high. In the fourth and fifth stages, Asus had owned profound knowledge and experiences in the China market and also found out the most optimum sales network pattern. It had lower transactions costs in fourth and fifth stages.

Asus' Channel Evolution

In the first stage of channel structure, Asus lacked domestic sales right. In order to quickly build up sales points, Asus had to cooperate with the local agents. Therefore, the channel structure on this stage is strategic alliance system. The purpose was to gain legitimacy. The difference from Acer was that their cooperating companies are their previous cooperated agents of main boards. At the second stage, Asus had domestic sales right. The institutional pressure became less. However, the agents of main boards were not good at selling notebook PC. Asus changed the strategy. It founded branch companies. The type of channel structure belonged to internationalization system. At the third stage, it considered entering the market but it must cooperate with local dealers. Thus Asus change to combined agents system. It means a combined operation system of chained shops and branch companies. The type of alliance system is to pursue legitimacy. Therefore it belongs to isomorphism alliance. At the fourth stage, for the reinforcement of its cooperative effects with Digital China, Asus established "Products Center Channel Department." Therefore the channel structure on the fourth stage belongs to alliance system in pursuit of efficiency. On its fifth stage, Asus cooperates with Highly Information Industry. Using the professional marketing capability of notebook PC industry of Highly Information Industry, Asus can

increase market share in China. Therefore, the channel structure in its fifth stage still belongs to alliance system in pursuit of efficiency.

CONCLUSIONS AND IMPLICATIONS

China domestic markets become larger and the competition become tougher day by day. Moreover, the life cycle of notebook PC is short and product homogeneity is very high. To avoid the price war, firms must perform well on channels management so that their products can get the market shares quickly. This research considers marketing channel as networking relationship and study the choices from internalization to externalization. We found that institutional environmental pressure becomes weaker and the transaction cost is relatively high, the Taiwanese firms' marketing channel tends to be internalization. As the institutional environmental pressure becomes weaker but the transaction cost is relatively low, the Taiwanese firms' marketing channel tends to be efficiency alliance.

REFERENCES

- [1] DiMaggio, P. J., Powell, W. L. W. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields," *American Sociological Review*, 1983, 48, 147 – 160.
- [2] Grewall, R., Dharwadkar, R. "The Role of the Institutional Environment in Marketing Channels," *Journal of Marketing*, 2002, 66(3), 82–97.
- [3] Heide, J. B. "Interorganizational Governance in Marketing Channels," *Journal of Marketing*, 1994, 58, 71–85.
- [4] Martinez, R. J., Dacin, T. M. "Efficiency Motives and Normative Forces: Combining Transaction Costs and Institutional Logic," *Journal Management*, 1999, 25(1), 75 – 96.
- [5] Oliver, C. "Strategic Response to Institutional Processes," *Academy of Management Review*, 1991, 16, 145 – 79.
- [6] Roberts, P. W., Greenwood, R. "Integrating Transaction Cost and Institutional Theories: Toward a Constrained-Efficiency Framework for Understanding Organizational Design Adoption," *Academy of Management Review*, 1997, 22(2), 246–373.
- [7] Suchman, M. C. "Managing Legitimacy: Strategic and institutional Approaches," *Academy of Management Review*, 1995, 20, 571–610.
- [8] Tesfom, G., Lutz, C., Ghauri, P. "Comparing Export Marketing Channels: Developed Versus Developing Countries," *International Marketing Review*, 2004, 21(4/5), 409–422.
- [9] Williamson, O. E. *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press, 1975.
- [10] Williamson, O. E. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. New York: The Free Press, 1985.
- [11] Williamson, O. E. "Comparative Economic Organization: The Analysis of Discrete Structural Alternative," *Administrative Science Quarterly*, 1991, 36, 269–296.

AN EMPIRICAL INVESTIGATION OF THE RELATIONSHIPS AMONG SERVICE QUALITY EXPECTATIONS, ACTUAL EXPERIENCES, AND ITS GAP TOWARD SATISFACTION

Arthur Cheng-hsui Chen, National Yunlin University of Science and Technology, Taiwan,

chencs@yuntech.edu.tw

Ya-hui Chang, National Yunlin University of Science and Technology, Tajen University, Taiwan,

ritachang23@yahoo.com.tw

Feng-Chun Fan, National Yunlin University of Science and Technology, Taiwan,

candy@mail.janfusun.com.tw

ABSTRACT

The objectives of this research are to examine: (1) the relationships among customers' quality expectations, quality perceptions of actual experiences, and satisfaction, (2) the gap between the expectation and actual experience toward customer satisfaction, (3) the impact of stating expectations before a consumption experience. A private theme park, located at the middle of Taiwan, was selected as the research object. High school students were recruited as research subject via convenient sampling when they visit the theme park. Two studies have been conducted. Study 1 focused on customers' quality expectations, actual experience and satisfaction. Study 2 contained consumers' perception of actual experience and their satisfaction only. Three hundred ninety two and four hundred forty four usable questionnaires were collected for both two studies respectively. The results showed that the overall customer expectation and actual experience perception have a significant positive relationship with customer satisfaction. Moreover, we find the mediating role of customers' perceptions of actual experiences on the relationship between expectation and satisfaction. On the other hand, the bigger gap between customer expectation and actual experience perception will lead to lower customer satisfaction. Finally, participants in the stated-expectations group evaluate perceptions of actual experiences higher than those in the without stated-expectations group. However, participants in the stated-expectations group evaluate satisfaction partially higher than those in the without stated-expectations group.

Key words: service quality, customer satisfaction, quality gap.

INTRODUCTION

Service quality has been recognized by the academicians and practitioners as one of the key concept of any organizations. Most of the research findings showed that service quality not only has a positive effect on customer satisfaction (Caruana, 2002; Cronin & Taylor, 1992; Spreng & Chiou, 2002; Spreng & Mckoy, 1996; Woodside, Frey & Daly, 1989), but also has a direct relationship with customer retention rates and company profitability (Backman & Veldkamp, 1995; Baker & Crompton, 2000; Bloemer, Ko de Ruyter & Wetzels, 1999; Zeithaml & Bitner, 2003). Among the literature, the disconfirmation model has become the dominant concept on service quality since the early 1980s. However, most of the researches of the service quality are focusing on the customer perception of retrospect, but less on customers' expectations or actual experience. Moreover, there is even less investigation on the gap between the expectation and actual experience toward customer satisfaction.

On the other hand, customer expectations have been consistently acknowledged as the base on which service quality and customer satisfaction judgments are formed (Oliver, 1993; Parasuraman et al., 1988). As we know, the important role of expectations has been proved in managing service delivery, and creating and sustaining long-term customer relationships (Boulding et al., 1993; Hamer et al., 1999; Kalamas et al., 2002; Ojasalo, 2001; Pitt and Jeantrout, 1994; Walker and Baker, 2000; Zabava Ford, 2001). Therefore, it is critical for us to find out about customers' expectations in advance when we examine service quality. However, the impact of stating expectations before a consumption experience and their subsequent evaluations of that experience is ambiguous. We are not sure if the higher expectation will lead to higher or lower customer satisfaction after actual experience.

Finally, much has been said about the relationships between service quality and customer satisfaction, but there is no empirical evidence to support it in outdoor recreation. As we know, the outdoor recreation attracts many young people to experience it as one of the key leisure activities. However, the perception and expectation of service quality in outdoor recreation may be different between adolescents and adults. Therefore, it motivates us to choose the context of outdoor recreation to examine the above research questions in service quality topic for adolescents.

With that in mind, the objectives of this research are to examine: (1) the relationships among customers' quality expectations, quality perceptions of actual experiences, and satisfaction, (2) the gap between the expectation and actual experience toward customer satisfaction, (3) the impact of stating expectations before a consumption experience.

LITERATURE AND HYPOTHESIS

The relationships among service quality expectations, actual experiences and satisfaction

Exceeding customer expectation is the key to customer satisfaction (Kotler, 2000). Therefore, it is important for marketers to find out about their customers' expectations in advance. Customer expectations are beliefs about a service that serve as standards against which service performance is judged (Zeithaml et al., 1993). Expectations are formed from a variety of sources, such as the customer's needs and wishes (Edvardsson et al., 1994), the customer's feeling about a service or product, by promise (staff, advertising, and certain service promise), by word-of-mouth communication, as well as by past experience of that service or product (Zeithaml and Bitner, 1996). In the theme park, for example, visitors are satisfied when the variety of recreation facilities meets their expectations. Moreover, they have the chance to take the recreation facilities. If the taking experience is excellent, visitor would satisfy with the trip on the park. Pine II & Gilmore said (1998) "as goods and services become commoditized, the customer experiences that companies create will matter most. Today the concept of selling experiences is spreading beyond theaters and theme parks." It is obviously that visitors' expectations would affect satisfaction through experiences. Thus,

H₁: Customers' service quality expectations have positive effect on customer satisfaction.

H₂: Customers' service quality expectations have positive effect on customers' perceptions of actual experiences.

H₃: Customers' quality perceptions of actual experiences mediate the effect of customers' quality expectations on satisfaction.

The relationship between the gap and satisfaction

The gap between expectations and actual experiences will have effect on satisfaction. When customers' expectations are higher than the actual experiences, there are negative perceptions and when customers'

expectations are lower than the actual experiences, there are positive perceptions. Customers who have the negative perceptions would evaluate the satisfaction negatively and those who have the positive perceptions would evaluate the satisfaction positively. Thus,

H₄: The gap between the expectations and actual experiences has positive effect on customer satisfaction.

H₅: Participants in the negative perceptions group evaluate the satisfaction lower than those in the positive perceptions group.

The effect of stating expectations

Prior research found that the need for customers to state expectations before a consumption experience is likely to make them more accessible during the experience (Feldman and Lynch, 1988). It also indicates that though stated expectations may be influenced by customers' preexisting expectations and information, they are often clarified and constructed when the need to mention them (Bettman, Luce, and Payne, 1998; Schwarz and Bohner, 2001). However, those findings do not lead to predict the impact of stating expectations on the valence of post experience evaluations.

The contrast theory predicts that when expectations are not matched by actual performance, the contrast between expectations and actuality will cause customers to exaggerate or magnify the incongruity. Thus, customers who state positive expectations tend to evaluate more negatively after the experience. Ofir and Simonson (2001, 2007) suggested that stating expectations before consumption leads customers to focus on negative aspects of that experience and produces more negatively results on subsequent evaluations. Alternatively, from the assimilation theory and confirmation bias, positive or negative expectations stated before a consumption experience largely reinforce prior beliefs and make more extreme post experience evaluations. The assimilation theory suggested that disconfirmed expectancy about a service or product will create psychological discomfort. As a result, consumer will reduce psychological tension by modifying assessment of that service or product. Dholakia and Morwitz (2002) suggest that stating expectations could have positive effect on subsequent evaluations when the measurement of expectations is equivalent to the measurement of satisfaction with past performance. Thus,

H₆: Participants in stating expectations group evaluate actual experiences higher than those in without stating expectations group.

H₇: Participants in stating expectations group evaluate satisfaction higher than those in without stating expectations group.

RESEARCH DESIGN

Sample and Data Collection

One of the most popular theme park with abundant world-class facilities in Taiwan was chosen as the research object. High school students were recruited as research subject via convenient sampling when they visit the theme park. Two studies have been conducted. Study 1 focused on customers' quality expectations, actual experience and satisfaction. The data were collected in two stages: the first one before customers' participation in the program in order to measure their quality expectations, and the second one after customers' participation in the program, in order to measure their actual quality perceptions. Study 2 contained consumers' perception of actual experience and their satisfaction only. Three hundred ninety two and four hundred forty four usable questionnaires were collected for both two studies respectively.

Measures of Constructs

Based on prior studies (Graefe et al., 2003) and adjusted to the context of the theme park, are used to measure the four dimensions of service quality expectations and perceptions of actual experiences. The four dimensions are as follows: services, facilities, information and recreation experience. We measure services with seven items that ask respondents, for example, to evaluate staff about services they provide. We measure facilities with five items that addressed the situation of related recreation facilities. We measure information with six items that reflected theme park's environment. Finally, we use six items to measure recreation experience, such as the respondents' feeling about the theme park. Customer satisfaction is measured with a 5-item scale taken from the related characteristics of the theme park, such as follows: how about the entire space arrangement, layout planning, theme park facilities, food and beverage facilities and employee services. A five point Likert-type scale ranging from "strongly agree" (5) to "strongly disagree" (1) is used.

ANALYSIS AND RESULTS

We use regression analysis to test the direct effects of customers' quality expectations and perceptions of actual experiences on customers' satisfaction. As Table 1 (Model 1) shows, customers' quality expectations on services ($b = .25$, $t = 5.58$, $p < .001$), facilities ($b = .24$, $t = 5.41$, $p < .001$), information ($b = .26$, $t = 5.89$, $p < .001$), recreation experience ($b = .20$, $t = 4.49$, $p < .001$) have a positive and significant effect on customers' perceptions of actual experiences. The result in Model 2 shows that customers' quality expectations on services ($b = .18$, $t = 3.84$, $p < .001$), facilities ($b = .17$, $t = 3.36$, $p < .001$), information ($b = .19$, $t = 3.96$, $p < .001$), recreation experience ($b = .16$, $t = 3.36$, $p < .001$) also have a positive and significant effect on customers' satisfaction. These results support H_1 and H_2 . When the four dimensions of customers' perceptions of actual experiences are entered into Model 3, they show positive and significant effect on customers' satisfaction. However, the four dimensions of customers' quality expectations become insignificant on satisfaction. According to the three-step regression procedure that Baron and Kenny (1986) recommend, it proves the mediating role of customers' perceptions of actual experiences. H_3 is supported.

The Fourth model, used to examine the gap between the expectation and actual experience toward customer satisfaction, shows that the gap on services ($b = .11$, $t = 2.2$, $p < .05$), facilities ($b = .17$, $t = 3.48$, $p < .005$), information ($b = .15$, $t = 3.08$, $p < .001$), recreation experience ($b = .19$, $t = 3.86$, $p < .001$) has a positive and significant effect on customers' satisfaction. Thus, H_4 is supported. In order to examine the effect of positive/negative perceptions on satisfaction, we use MANOVA analysis. Participants in the negative perceptions group (Group 1, $N = 202$) evaluate the satisfaction lower than those in the positive perceptions group (Group 2, $N = 190$): For the space measure, $MG1 = 3.40$ versus $MG2 = 3.67$ ($F = 12.44$, $p < .001$); for the layout measure, $MG1 = 3.53$ versus $MG2 = 3.74$ ($F = 6.85$, $p < .05$); for the facilities measure, $MG1 = 3.54$ versus $MG2 = 3.84$ ($F = 12.54$, $p < .001$); for the F&B facilities measure, $MG1 = 3.04$ versus $MG2 = 3.37$ ($F = 8.69$, $p < .001$); for the services measure, $MG1 = 3.37$ versus $MG2 = 3.76$ ($F = 20.54$, $p < .001$). These results support H_5 .

Finally, Model 5 shows customers' quality perception of actual experiences without stating expectations on satisfaction. The actual experiences without stating expectations on services ($b = .31$, $t = 8.81$, $p < .001$), facilities ($b = .27$, $t = 7.63$, $p < .001$), information ($b = .42$, $t = 11.8$, $p < .001$), recreation experience ($b = .32$, $t = 8.99$, $p < .001$) have a positive and significant effect on customers' satisfaction. MANOVA analysis is used to examine the impact of stating expectations before a consumption experience. After visiting the theme park, participants in the stated-expectations group (Group 1, $N = 392$) evaluate perceptions of actual experiences higher than those in the without stated-expectations

group (Group 2, N = 444): For the services measure, MG1 = 3.72 versus MG2 = 3.55 (F = 12.17, $p < .001$); for the facilities measure, MG1 = 3.82 versus MG2 = 3.65 (F = 13.55, $p < .001$); for the information measure, MG1 = 3.73 versus MG2 = 3.62 (F = 5.67, $p < .05$); for the recreation experience measure, MG1 = 3.79 versus MG2 = 3.60 (F = 18.71, $p < .001$) (Table 3). These results support H6. As Table 4 shows, after visiting the theme park, participants in the stated-expectations group (Group 1, N = 392) evaluate the satisfaction higher than those in the without stated-expectations group (Group 2, N = 444): For the facilities measure, MG1 = 3.69 versus MG2 = 3.52 (F = 9.03, $p < .001$); for the F&B facilities measure, MG1 = 3.20 versus MG2 = 3.04 (F = 5.12, $p < .05$); for the services measure, MG1 = 3.56 versus MG2 = 3.30 (F = 18.53, $p < .001$). However, there is no significant difference evaluation on satisfaction for the space measure (MG1 = 3.53 versus MG2 = 3.52 (F = 0.06, $p = .80$)) and the layout measure (MG1 = 3.63 versus MG2 = 3.53 (F = 3.55, $p = .06$)). Thus, H7 is partial supported.

DISCUSSIONS AND CONCLUSIONS

This study is to investigate of the relationships among service quality expectations, actual experiences, and its gap toward satisfaction. It builds on previous service quality research conducted by both consumer behavior and recreation researchers in an attempt to better understand customer satisfaction. In congruence with previous research on satisfaction, many of the constructs associated with quality in an outdoor recreation environment are services, facilities, information, and recreation experience.

The hypothesized relationships among service quality expectations, actual experiences and its gap toward satisfaction are upheld. Customers' service quality expectations directly affect both actual experiences and satisfaction. When customers' service quality expectations and actual experiences are taken into considering on satisfaction, we find the mediating role of customers' perceptions of actual experiences. Outdoor recreation is a unique and highly intangible activity that relies on the provision of quality services from recreation providers. Visitors' recreation experience is also the multi-dimensional, transitory and multi-phased process (Lee, Dattilo & Howard, 1994). We conduct the experiences measure immediately after participants finishing their trip on the park. Outdoor recreation, an intangible activity, becomes a more tangible activity through actual experiences. By that way, actual experiences reinforce the customers' satisfaction.

Moreover, the gap between the expectations and actual experiences also affects customer satisfaction. Participants whose quality expectations are higher than actual experiences have the negative perceptions and participants whose quality expectations are lower than actual experiences have the positive perceptions. The different perceptions result in different evaluation about satisfaction. Visitors in the negative perceptions group evaluate the satisfaction lower than those in the positive perceptions group. Both groups evaluate the satisfaction on Food and Beverage facilities domain lower than that on the other four domains. For the theme park managers, they should pay more attention to this aspect in order to maintain consistently quality service.

Generally speaking, most researchers acknowledge that customers have expectations and that they plan an important role as standard or reference points used by consumers to evaluate the performance of a company (Oliver, 1980; Brown and Swartz, 1989; Bolton and Drew, 1991a, 1991b; Zeithaml et al., 1990; Parasuraman et al., 1985, 1988, 1994). In this study, actual experiences plan another important role when customers evaluate the performance of a company. As Burns et al. (2003) found, a visitor's recreation experience is a key issue for recreation managers to evaluate. In order to obtain customer

satisfaction, managers should understand customers' expectations first and then create an exceeding expectations environment.

Another concern with the impact of stating expectations before a consumption experience is proven positively on customers' evaluation of actual experiences. Participants in the stated-expectations group evaluate perceptions of actual experiences higher than those in the without stated-expectations group. However, it is proven positively on customers' evaluation of satisfaction on three domains (facilities, F&B facilities and service). In our study, the theme park is famous at its abundant world-class facilities and is popular among adolescents. Adolescents visit the park with peers to have exciting experience and may only notice the attributes they experienced. When adolescents evaluate about satisfaction, they would be satisfied with recreation and F&B facilities they met and service they feel.

Table 1 Results of Regression Analysis: Standardized Coefficients (t-Values)

Independent Variables	Actual experience	Satisfaction			
	Model 1	Model 2	Model 3	Model 4	Model 5
Expectation					
Services	.25(5.58)***	.18(3.84)***	.04(.89)		
Facilities	.24(5.41)***	.17(3.63)***	.01(.29)		
Information	.26(5.89)***	.19(3.96)***	.02(.46)		
Recreation experience	.20(4.49)***	.16(3.36)***	.04(.87)		
Actual experience					
Services			.29(7.09)***		
Facilities			.34(7.94)***		
Information			.34(7.98)***		
Recreation experience			.28(6.97)***		
The gap					
Services				.11(2.2)*	
Facilities				.17(3.48)**	
Information				.15(3.08)**	
Recreation experience				.19(3.86)***	
Actual experience (without stating expectations)					
Services					.31(8.81)***
Facilities					.27(7.63)***
Information					.42(11.80)***
Recreation experience					.32(8.99)***
F value	28.79	13.72	35	10.34	88.96
Adjusted R2	0.22	0.12	0.41	0.09	0.44
N	392	392	392	392	444

***p < .001, ** p < .005, * p < .05

References is available on request

IMPACTS OF FINANCIAL CRISIS AND ITS THEORETICAL IMPLICATIONS

H. C. Li, Bryant University, Smithfield, RI 02917
hli@bryant.edu, (401) 232-6079
Victor Lux Tonn, Salve Regina University, Newport, RI 02840
jpmaple@aol.com, (401) 848-5821

ABSTRACT

The current financial crisis has its microeconomic as well as macroeconomic causes. At the microeconomic level, the root of the problem can be traced to consumer attitude towards debt, business psychology of financial innovation, the growth of institutions which create, market and buy mortgage backed securities, and the globalization of financial markets. At the macroeconomic level, the raising interest rates to combat inflation have also contributed to the burst of the housing bubble.

To capture the interaction between the financial sector and the real sector, we suggest the modification of the quantity theory of money. Instead of having a term defined as the product of price level and aggregate real product, we use the sum of price and quantity of individual products. In addition, the price and the quantity of each product are each weighted by a magnifying factor which captures the impact of volatilities of financial market.

Keywords: sub-prime mortgage crisis, financial sector, real sector, quantity theory of money.

CAUSES OF THE CURRENT FINANCIAL CRISIS

The current financial crisis has its microeconomic as well as macroeconomic causes. At the microeconomic level, the problem can be traced to the historical roots of financial innovations and institutional changes of the mortgage industry.

The U. S. mortgage market has evolved drastically since the New Deal. Originally, mortgage lending was dominated by local depository institutions regulated by government agencies. Secondary mortgage market virtually did not exist. Credit risk, interest risk and liquidity risk were all assumed by mortgage originators. The Great Depression led to the development of a secondary mortgage market. As a part of the New Deal, the Federal National Mortgage Association (FNMA) was created in 1938 to acquire portfolios of conforming mortgages from local lenders. Thus, originators were able to transfer their credit risk, interest risk and liquidity risk to a government agency. In subsequent years, FNMA was privatized. The Government National Mortgage Association (GNMA) and the Federal Home Loan Mortgage Association (Freddie Mac) were created. Through financial innovations in the forms of mortgage guarantees and securitization, these agencies have been able to channel additional capital into the mortgage market, and at the same time, to diversify the risk of mortgage lending [1].

These agencies have their activities limited basically to conforming or prime mortgages. Their success have however attracted private institutions to the secondary mortgage market and the growth of Alt-A and sub-prime mortgage products. The introduction of collateralized mortgage obligations and other mortgage derivatives has lured an even broader range of investors to trade mortgage related financial instruments. The growth of mortgage back securities, together with the lack of government regulation and supervision, have further promoted the development of Alt-A and sub-prime mortgages products such as NINJA (No Income, Job, or Asset) and interest only loans. Thus, financial innovation and risk diversification have sowed the seeds of future disaster.

In addition to profit motivation of financial innovation, the root of the problem can also traced to consumer attitude towards debt. It is well known that the propensity to save in the U.S. is lower than many other countries. The growth of consumer spending and debt is further fueled by the advancement of electronic technology. The use of credit cards allows consumers to obtain easy credit to finance their spending. Eventually, many households have been driven close to or into personal bankruptcy.

Traditional economic and financial theory is built on the foundation of rationality. Consumption and investment decisions are based on the concept of utility maximization. Thus, the behavior of simply following the lead of others has been ruled out by the mainstream economics and finance. Behavioral economics and finance, on the other hand, suggest that fad or herding is not irrational at all. The housing bubble and the strong demand for high risk mortgage backed securities can be viewed simply as a pattern of herding [3].

Financial innovations and institutional development function like double-edged swords. They were responsible for the growth of sub-prime mortgage market; they have also fueled and deepened the sub-prime mortgage crisis. The promotion of home equity loans by banking institutions, a profitability strategy before the crisis, has lowered the returns of lenders since the burst of the housing bubble. While credit lines extended to hedge funds and other speculators constituted an important component of the off-balance sheet activities of depository institutions, they then became a major source of the liquidity problem. The highly leveraged operating strategy, which brought handsome returns to hedge funds, forced many top executives of Wall Street power houses to step down. Some of the mortgage companies which were proud of their innovative mortgage designs, laid off thousands of their employees and even put themselves on the market for sale. The trouble in the mortgage market soon led to the credit crunch and disaster in commercial paper market. Globalization of financial markets contributed to the worldwide risk diversification of mortgage-backed securities, it also quickly spread disaster overseas. Investors failed to accurately evaluate the quality and impacts of many innovative products and performances of financial institutions; so did the rating agencies and the regulators.

At the macroeconomic level, the threat of high rate of inflation led to the rise of interest rates, which in turn dealt a severe blow to the housing industry and the mortgage market. That is, steady increase in interest rates lowered the housing demand. At the same time, many of those who borrowed adjustable-rate mortgages failed to make payments on schedule. As a result, mortgage backed securities suffered large decrease in prices. The impact was not limited to the housing and mortgage industries. Industries such as home furnishing and appliances also felt the

pain. Through the multiplier effect, the whole economy slowed down. Globalization of financial markets and free trade agreements immediately made the crisis a worldwide economic problem.

To cope with the sub-prime crisis, attempts were made by the Federal Reserve and other its counterparts in other countries to infuse liquidity into their respective financial systems. Central banks also lowered interest rates for preventing the occurrence of a national or worldwide recession. Efforts of the monetary authorities have been accompanied by other government actions, including changes in regulations, actions of government sponsored agencies, and changes in fiscal policies.

THEORETICAL IMPLICATIONS

Numerous attempts have been made by economists to construct a microeconomic foundation of macroeconomics. Up to this date, no widely acceptable theoretical model has been introduced. Separately, financial sector of an economy is often treated as an image of the real sector. We suggest that the formation and the burst of a financial or speculative bubble are closely related to the interaction between the financial sector and the real sector, as well as the interaction between microeconomics and macroeconomics. Fluctuations in the real sector could lead to volatility of the financial market. In turn, cascades of investment returns further intensify the cyclical movements of the real sector. Changes in interest rates have significant consequences on the housing sector, while the burst of the housing bubble results in the slow down of the whole economy. The interaction between finance and real sectors and between individual industries and the whole system, through the multiplier effect, result in cyclical movements of the economy.

To capture the interaction of between the financial sector and the real sector, and between microeconomic and macroeconomic activities, we suggest the modification of the quantity theory of money. In a nutshell, the original quantity theory of money suggests that real output (Y) multiplied by its price level (P) is equivalent to the product of money supply (M) and velocity (V). That is,

$$MV = PY. \tag{1}$$

Instead of having the nominal aggregate output defined as the product of price level and aggregate real product, or PY, we use the sum of individual products, each multiplied by its corresponding price. In addition, the quantity of each product and price is weighted by a magnifying factor, which captures the interaction between the components of financial sector and the components of real sector. Thus, the modified quantity theory of money becomes

$$MV = \sum(n_i P_i)(m_i Y_i), \tag{2}$$

Where

M = money supply,

V = velocity,

i = the *i*th industry,

P_i = price of the *i*th product,

n_i = magnifying factor showing the impact of the interaction individual sectors on P_i ,

Y_i = output of the i th industry, and
 m_i = magnifying factor showing the impact of the interaction individual sectors on Y_i ,

Note that the price of a financial industry is the interest factor, $(1 + r)$, where r is the interest rate or rate of return. In this model, at least one of the industries belongs to the financial sector.

Any change in Y_i , given the values of other components on the right-hand side, triggers a change in V , as the impact of the real sector reaches the financial sector. An adjustment in V , a reflection of fluctuations in the financial market, however, is expected to have a feedback effect on the real sector through the change in n_i and m_i . It is worth to mention that, even when M and V are held constant, price of quantity change in one industry could be triggered by changes in another industry. As a result, the nominal value of aggregate output, $\sum(n_i P_i)(m_i Y_i)$, is affected by changes in individual industries as well as changes in M and V .

The current financial crisis provides an excellent opportunity to examine the interaction between microeconomic units and their corresponding macroeconomic system. This interaction in the real sector is amplified by the crisis in the financial sector. Up to this point, the sub-prime mortgage crisis has its microeconomic as well as macroeconomic causes. The interaction between these sectors, through the magnification process, result in the slow down of the whole system. The weakening of the economy, in turn, prevents the housing sector and the mortgage market from the quick recovery.

Successful estimation of n_i and m_i would provide a better understanding of the linkage between microeconomics and macroeconomics, and to better explain the relationship between the real sector and the financial sector. It is known that returns on many and most financial instruments are not normally distributed. This phenomenon is often attributable to the herding behavior of investors [2]. Similarly, sharply changes in price of a product can also be explained by the formation of fad. The testing of the suggested model requires new set of econometric tools.

REFERENCES

- [1] Dodd, R. "Subprime: Tentacles of a Crisis," *Finance and Development*, 2007, 44 (4), 15-19.
- [2] Mantegna, R., Stanley, H. E., *An Introduction to Econophysics, Correlations and Complexity in Finance*, Cambridge, U.K.: Cambridge University Press, 2000.
- [3] Nofsinger, J. R. *The Psychology of Investing*. Upper Saddle River: Pearson Education, 2002.

**GOVERNMENT'S ROLE IN THE REFORM OF BANKING SYSTEM:
EVIDENCE FROM CHINESE STATE-OWNED COMMERCIAL BANKS**

Qinghua Song

Xinhua School of Finance and Insurance, Zhongnan University of Economics and Law,
Phone: 86-27-6271-1762 Email: songqh917@gmail.com

Chung-Jen Fu

Department of Accounting, National Yunlin University of Science & Technology
Phone: 886-5-534-2601#5341 Email: fucj@yuntech.edu.tw

Bing-Xuan Lin

College of Business Administration, University of Rhode Island
Phone: 401-874-4895, Email blin@uri.edu

ABSTRACT

We review the commercial banking system reform in China and show that the approaches taken by the Chinese government have been quite successful. Using multiple mechanisms such as capital injection, disposal of non-performing loans, preferential interest rate, tax incentives and foreign strategic partners, Chinese government has successfully converted three major banks in China from state-owned enterprises to market-oriented stock companies. Our estimation shows that the equity interests held by the government in these three banks have greatly outweighed the cost of restructuring and the loss incurred in restructuring these banks. Although Chinese banks still lack experience in corporate governance, risk management and product innovation, they have presented a unique approach of banking system reform.

Keywords: Chinese banking system; Bank reform; Non-performing loans; Role of government

Chinese financial system is largely a bank-driven system. The four major banks: Industrial and Commercial Bank of China (ICBC), Agricultural Bank of China (ABC), Bank of China (BOC) and China Construction Bank (CCB), hold more than 50% of the national deposits and over 50% of the loan business. These four banks have played instrumental roles in the growth of Chinese economy; however they had also large amounts of bad debts and experienced severe operation inefficiency. Around 1990s, all four banks were on the verge of bankruptcy (Song,2000). Banking system reform became inevitable in order to secure the market stability. Starting from 1994, Chinese government started to reorganize these banks through financial restructurings and capital market operations. The highlights of the banking system reform have been the successful IPOs of three major banks. The experience of ICBC, BOC and CCB suggests three interesting solutions to banking system reform: (1) effective government intervention, (2) successful use of capital markets and (3) adoption of bank-specific approach.

The most interesting aspect of the reform is the unique approaches applied to different banks. Three of the four major commercial banks are now publicly traded corporations. Specifically, BOC and ICBC were transformed into Corporations in its entirety. CCB was broken into one banking entity (China Construction Bank Corporation) and one non-banking entity (China Jianyin Investment Corporation). The reform of ICBC was co-sponsored by the Ministry of Finance and the government investment company called Central Huijin Investment Corporation (Huijin). BOC is solely sponsored by Huijin. The marketization of CCB involves five entities including Huijin, Jianyin and etc. Between 2004 and 2005, Huijin injected over \$60 billion in the process of restructuring the above three banks. Furthermore, banks also formed strategic partnership with various foreign banks in U.S., Europe and Asia. Finally, the process of IPO also differed for these banks. ICBC cross-listed in Shanghai and Hong Kong simultaneously; BOC and CCB were initially traded in Hong Kong and then returned to mainland China.

In this article, we review the reform of the four major State-owned commercial banks and examine the government's role in the banking system reform. Part II of the paper describes how Chinese government restructured these banks through government intervention. Part III evaluates the effectiveness of China banking system reform from the perspectives of the government and the banks. Finally, Part IV concludes with our final thoughts and some policy recommendations.

II. Government's Role in the Banking System Reform

1. Capital Injection

Prior to market reform, all state-owned commercial banks in China were solely owned by the Chinese government and initial capital was funded by the government as well. Meanwhile, all profits and losses were seized by the government. Due to limited funding channel and poor operation performance, these commercial banks traditionally had very low capital adequacy. Starting from 1998, the government had injected capital both directly and indirectly into these banks. In 1998, MOF issued 270 billion Yuan treasury notes and the proceeds were used to augment the capital reserve of these banks. In 2003, the government established Central Huijin Investment Corporation based on the investment of the foreign exchange reserves by State Administration of Foreign Exchange. Huijin becomes the most important vehicle for the government to inject new capitals in state-owned commercial banks. Two weeks after its origin, Huijin injected \$45 billion in BOC and CCB. In 2005, Huijin provided \$15 billion for ICBC. The government also encouraged these banks to raise capital through issuing new stocks, bonds and establishing strategic alliances with foreign banks. In 2005, CCB went public on Hong Kong Stock Exchange and raised HK\$71.58 billion through IPO. Two years later, CCB raised another 57.1 billion Yuan through an IPO offer in Shanghai Stock Exchange. In 2006, BOC raised 20 billion Yuan in mainland China and HK\$86.7 in Hong Kong. Finally, ICBC offer both A and H shares on Oct 27th, 2006 and became the largest IPO globally. It raised HK\$124.9 in Hong Kong and 46.4 billion Yuan in mainland China. In addition to equity offerings, government also announced several guidelines to facilitate the securitization of sub-prime long-term loans held by these banks. By the end of 2006, the debt issued by BOC, CCB and ICBC approximated 16.28

billion Yuan. The last source of capital came from foreign strategic investors. Chinese government approved a list of Foreign Institutions and provided great incentives for them to work with Chinese banks. Meanwhile, each Chinese bank was encouraged to seek foreign partners to boost its capital reserve. For example, BOC teamed up with Bank of Scotland, CCB partnered with Bank of America and ICBC aligned with Goldman Sach. The investment of foreign partners totaled more than 106 billion Yuan. The details of funding amount are reported in Table 1 by bank and by sources. We can see that capital increased by over 1422 billion Yuan. For three listed banks, the amount of capital injection exceeded 400 billion for each bank.

Table 1 Capital Injection through Different Sources in billions of Yuan

	Treasury Note	Huijin	Foreign Partners	IPO: Shanghai	IPO:Hong Kong	Debt Issuing	Total
ICBC	85	124.0	30.5	46.6	126.6	35	447.7
BOC	42.5	186.4	43	20	90	60	441.9
CCB	49.2	186.2	32.8	57.1	74.6	40	439.9
ABC	93.3						93.3
Total	270	496.6	106.3	123.7	291.2	135	1422.8

Source: Okazaki (2007) and Banks' Annual Reports

2. Disposal of Non-Performing Loans (NPLs)

One of the overhanging problems for State-owned commercial banks was the colossal amount of non-performing loans. In 1999, the ratio on NPL to total loans for State-owned commercial banks was up to 40% and it was still over 20% (\$232 billion) in 2003(Yuan and Zhao,2007). Effectively, there were three rounds of NPL disposals since 1994. The first round took place in 1994 when State-owned commercial banks disposed their NPLs to three newly formed policy banks: China Development Bank, Export-Import Bank of China and Agricultural Development Bank of China. For instance, by the end of 1994, ICBC, ABC and BOC disposed over 356 billion Yuan of loans to the Agricultural Development Bank of China. The second round of disposal happened in 1999 and four asset management companies (Orient Asset Management, Great Wall Asset Management, Cynda Asset Management, Huarong Asset Management) were created to absorb the NPLs for different banks. All together, the four major banks disposed over 1,294 billion Yuan of NPLs to the four asset management companies. By the end of 2006, the above four asset management companies unloaded 866.34 billion Yuan of NPLs and collected 180.56 billion Yuan. After the disposal of NPLs to asset management companies, the balance sheet quality for banks improved greatly. However, these banks failed to improve their operating efficiency and new loan quality, NPLs rose again to the original level just after a few years. The final round of NPL disposal took place prior to the IPO of these commercial banks. The three banks that raised capital though public offerings disposed 1162.8 billion Yuan of NPLs to government sponsored asset management companies. Clearly, Chinese government played a

crucial role in disposing the NPLs of its major commercial banks. The first round of disposal mainly dealt with NPLs created by the government economic policy prior to its adoption of market economy. It is only fair for the government to step in and absorb those NPLs. However, the government should not be responsible for the increasing NPLs completely in the second and third round of disposal. Operation inefficiency, low management quality, incompetent personnel, weak protection of debt-holders and etc have all contributed to the drastic increase of NPLs (Zhou,2005). In order for these banks to go public with a clean slate, the government stepped in and assumed more than the fair share of their responsibility.

3. Policy Support

The banking system reform was led by the Chinese government and supported by various government agencies and regulatory bodies. In 2002, State Council of the People's Republic of China formed a team to lead the reform of State-owned commercial banks. In 2003, the decision was made to convert exiting banks into publicly traded corporations. Many policies and rulings were made, and new agencies (including Huijin and AMCs) formed to support banks' effort in raising capital reserve and lowering NPLs. In 2007, National People's Congress approved the proposal of issuing 1,550 billion Yuan Treasury notes to establish China Investment Corporation(CIC) that became the parent company of Huijin. From the beginning of the banking system reform to the IPOs of the major banks, the government had created favorable policies and provided various incentives to facilitate the transformation of its financial system. The government also provided direct support through its tax policy and interest rate policy. In 1997, the government lowered the income tax rate for all State-owned commercial banks from 55% to 33%. However the business tax was raised from 5% to 8%. In 2001, the government began to take steps to reduce the business tax, and it returned to 5% in 2004. A recent proposal was to further reduce the business tax to 2% and eventually removing it (Tian,2006). Tax reduction was also awarded to major commercial banks. For example, CCB received a tax reduction of 15.74 billion and 7.85 billion in 2004 and the first half of 2005.

Another effective policy support was the interest rate policy. For a long period of time, commercial banks in China were only allowed to lend and borrow using the government approved rate. The requirement was then relaxed to allow lending rate to float within the range allowed by the central bank. Starting from October 2004, commercial banks in China were allowed to lend at 10% below the published rate and the ceiling for the lending rate was removed. The difference between the published one-year deposit and lending rates has been kept above 3.33% after 1999. This is relatively larger than the rates in developed countries and it helps to ensure a higher profit margin for commercial banks in China. The favorable interest rate policy ensured net interest margin and the ratio of net interest revenue for BOC, CBC, ICBC to maintain between 2% and 3% between 2003 and 2007.

III. Effectiveness of Chinese Banking System Reform

1. Positive results for the Government

The Chinese government has directly and indirectly funded the banking reform through many rounds of capital injection, NPL disposals and a series of favorable policies. However, the government has maintained absolute control of these publicly traded commercial banks after their IPO. If we examine the holdings of the government in the three banks as shown in Table 2, we see that the total market valuation of the holdings adds up to 4509.48 billion Yuan (3952 billion Yuan above the book value).

Table 2 Ownership of Listed Banks by the Chinese Government

Bank	Institution	Shares (billion Yuan)	Book Value (Yuan/Share)	Total Cost (billion Yuan)	Market Price (Yuan/share)	Market Cap (billion Yuan)	Premium (billion Yuan)
ICBC	MOF	124	1.00	124	8.13	1008.12	884.12
	Huijin	124	1.00	124	8.13	1008.12	884.12
BOC	Huijin	171.325	1.00	171.325	6.61	1132.46	961.13
CBC	Huijin	138.15	1.00	138.15	9.85	1360.78	1222.63
Total		557.475		557.475		4509.48	3952

Source: IPO prospectus. Market Capitalization is based on the 2007 last trading day price of A share.

The expense incurred by the government mainly includes the following: 1) 270 billion Yuan capital injection in 1998; 2) investment of \$60 billion (496.6 billion Yuan) by Huijin; 3) cost of disposing 2456.8 billion Yuan of NPLs (conservatively we estimate the cost might run up to 2000 billion Yuan). The total cost adds up around 2,770 billion Yuan, 1,800 billion Yuan less than the total premium in the value share ownership. The simple calculation (ignoring indirect holdings of Huijin) suggests that the government have earned significant premium in the process of banking reform. However, it is worth noting that the positive outcome relies heavily on the favorable market price of these banks.

2. Improved Performance for Listed Banks

Two important outcomes of the banking system reform are the higher capital adequacy ratio and lower NPL ratio. Prior to 1998, none of the four major banks met the minimum capital adequacy ratio of 8%. After the reform, the capital adequacy ratio for three listed banks all surpassed 12% in 2006. The NPL also dropped significantly for all three listed banks. For example NPL ratio for CCB went from 16.97% in 2003 to 2.95% in June 2007, and NPLs for BOC and ICBC were 3.56% and 3.29% in June 2007. Meanwhile, the listed banks also learned from their foreign partners and went through organizational changes to manage risk exposure, optimize operation, offer new products and improve efficiency. Tools such as risk adjusted return on capital and economic value added were applied to measure bank performance. Return on assets (ROA) and return on equity (ROE) also improved for listed banks. Due to the financial restructuring and the policy support, three listed banks earned 11-14% ROE in the third quarter of 2007. Table 3 presents the performance comparison among four major banks. We can see that three listed banks experienced a higher ROA and ROE after they became stock companies.

Name	Performance	2003	2004	2005	2006	2007
ICBC	ROA	0.49	0.61	0.59	0.66	0.74
	ROE	-4.20	-6.17	14.51	10.47	12.14
BOC	ROA	0.80	0.59	0.69	0.89	0.83
	ROE	13.90	10.85	12.41	11.29	11.10
CBC	ROA	0.01	1.25	1.03	0.85	0.89
	ROE	0.22	25.08	16.37	14.03	14.64
ABC	ROA	0.06	0.05	0.02	0.11	-
	ROE	1.39	2.57	1.31	6.91	-

Source: Wind Database and Bankscope.

IV. Conclusions

Generally speaking, the banking system reform in China has been successful and all participants have reached the win-win outcome. The experience for the three listed commercial banks can serve as models for other banks ready to follow suit. ABC, joint-stock commercial banks, urban commercial banks, rural credit unions and many other financial institutions can derive their unique approaches and use capital markets to achieve their goals. However, it is worth noting that government control remains strong even for listed banks. To improve corporate governance and operation efficiency, the government needs to gradually relax its control and allow the market to be the dominant force in the banking system. Huijin and government sponsored investment companies have earned significant premium due to favorable market conditions, but shares held by these institutions are still in the lock-up period and are not freely traded yet. The uncertainty in future stock market can have a huge impact on the outcome of the ongoing banking system reform. As for commercial banks that have successfully gone through the IPO process, there is no guarantee that future performance will continue to improve (García-Herrero, Gavilá and Santabárbara, 2005). Moving from a government-sponsored banking system to a market-driven operation, banks need to go through significant changes internally. Globalization has provided Chinese banks great opportunities to learn from their strategic partners; meanwhile it also poses a great challenge for Chinese banks to compete with foreign banks that are more experienced and technologically sophisticated. The banking system reform is not yet complete and Chinese government needs to vigilantly monitor the economy and the banking system to meet any new challenges.

References available upon request from Qinghua Song

CAPE ATLANTIC SAVINGS BANK: TRANSFORMING INTERNET BANKING

Haidong Liu, Bloomsburg University, (570)389-4564 hliu2@bloomu.edu
David Krause, Richard Stockton College, stk27675@loki.stockton.edu

ABSTRACT

This case concerns Cape Atlantic Savings Bank's (CASB) dilemma on how to transform the Internet Banking system into a full service distribution channel. The first step is to analyze the current Internet banking system along with its associated costs and impact of the institution. Secondly, the Bank must identify the enhancements to the current system that are needed to transform this Internet Banking system into a full service distribution channel. Finally, enhancements available from the current Internet Banking service provider are reviewed as well as those available from alternate service providers.

CASE SYNOPSIS

The Information Technology committee of CASB met early one Friday morning in January. Vice President of Information Technology Preston Smith informed the committee that the Bank's contract with Internet Banking service provider, Digital Insight, was set to expire in August. A discussion among committee members ensued. Senior Vice President Juliana Hatfield stated that she felt the Bank was not realizing the full potential of the Internet Banking channel. She added that ever since its implementation in 2001, the system has been a loss leader for the Bank. Assistant Vice President of Information Technology Blair Betts informed the committee that there have been many enhancements to Internet Banking systems since 2001 that would enable Internet Banking to become a truly virtual branch of the Bank. Executive Vice President Jason Ward stated that now was the time to review the Internet Banking channel and evaluate its cost and effectiveness to determine the impact the system has had on the Bank. He ordered that in addition, alternate vendors should be researched to ensure the Bank is not only getting the best product on the market, but one that is competitively priced.

COMPANY BACKGROUND

CASB is an eight branch federally chartered savings bank operating in New York. The Bank's main areas of business include retail/branch banking, residential mortgage lending, consumer lending and small business lending. Founded in 1887, one of the Bank's top goals has been to develop a strong, long-lasting relationship with the customer by continually offering innovative products and services. CASB has used superior customer service to create a niche for itself in the jungle of large commercial banks. Customers enjoy the personalized service they receive at CASB and enjoy returning to a Bank where the employees remember their name.

The Bank is guided strategically by President Kyle Reese and the Board of Directors, which is composed of six independent directors. In less than ten years as president, Mr. Reese has transformed CASB from a two branch state chartered savings and loans with less than \$100 million in assets to an eight branch federally chartered savings bank with more than \$500 million in assets and growing quickly.

Bank assets nearly doubled in the last two years. The Bank attributed the growth to the sales culture that was developed in 2001. Prior to 2001, the Bank relied on word of mouth and media advertising to attract new deposit and loan customers. The new sales program has changed the Bank's branch managers into sales people who solicit businesses and individuals in an effort to bring their business to CASB. The program has been quite successful in bringing in new deposits to the Bank and has helped to increase net income almost 100% in two years.

Customer deposits are invested in 1-4 family mortgages, mortgage-backed securities and government securities. The Bank's primary revenue was derived from the interest earned on residential mortgage loans originated by the Bank and interest earned on investments in mortgage-backed and government securities. A small portion of revenue was generated from the various fees charged to Bank customers. The Bank's two largest expense categories have been interest payments to depositors and employee salaries and related expenses.

THE INTERNET BANKING ENVIRONMENT AT CASB

CASB has always placed a premium on the use of technology to better service customers, achieve operating efficiencies and improve overall profitability. In 1998, CASB converted its core processing system from an IBM AS/400 mainframe system to client/server system architecture. The bank outsources its technology functions to Open Solutions, Inc. of Glastonbury, Connecticut using their Total Client Server core application as well as several supporting functions. Total Client Server is a comprehensive banking system that includes modules for savings, demand deposits, mortgages, commercial and consumer lending. Total Client Server uses an Oracle relational database that can be accessed by client workstations that have been configured on the Windows 2000 platform.

All of the data associated with the Bank's savings, checking and loan records are maintained on the Bank's data server, which is maintained by Open Solutions at their data processing center. The data server acts strictly as a database for the application system to query. The system operates in a UNIX/Oracle environment with a wide area network connecting all of the Bank's eight branches and departments into one network. The application systems (i.e. savings, demand deposit, mortgages, consumer, commercial, etc.) run on the client workstations such as the teller, customer service representative or back office personnel's personal computers. All transaction processing takes place on the client workstation. The clients retrieve and store data on the server at Open Solutions.

In 2001, CASB began to offer Internet Banking services to customers. The Bank selected Digital Insight, a business partner of Open Solutions, to provide the Internet Banking platform. As a business partner of Open Solutions, Open Solutions provided a real time interface between the Internet Banking system and the Total Client Server core processing system. The Digital Insight Internet Banking product offered customers the ability to view account history and balances in a real time environment as well as transfer funds between their accounts held at CASB. The system also enables users to view images of checks that have been paid against their accounts. At the inception of this system in 2001, CASB was the only bank in the area to offer online check images. That position did not last long as numerous banks followed suit.

The Digital Insight system also offered bill payment services through a business partnership with Metevante Corporation. CASB offered the Internet Banking product to customers for free and Bill Payment was offered to customers at \$5 per month. Within the first year, the \$5 monthly bill pay fee to customers was eliminated. The reasoning was twofold: 1) customer acceptance of

the product was low and 2) competitors were offering the bill pay service to customers for free. This has since become the industry standard.

In 2006, the Bank added Online Statements to its offering of online banking services. Recent changes to federal regulations allowed customers to opt out of receiving their paper statements. Initial customer acceptance of this product has been slow, but the Bank expects wider acceptance of this service by 2008.

Performance Assessment of the Current Internet Banking Environment

From its outset, Internet Banking has been a loss leader for CASB. The Information Technology Committee asked Internal Auditor Henry Jones to review the performance and costs associated with the Internet Banking system. Mr. Jones was able to obtain performance data back to 2001 for Internet Banking and 2003 for Bill Payment.

Mr. Jones' first task was to determine the cost per internet banking transactions. The expenditures were averaged over six years to account for capital expenditures and provide a more accurate overall cost per transaction. Internet Banking and bill pay transactions have grown at approximately 63% per year since inception. The cost per transactions has decreased significantly during this time frame. User acceptance was slow during the first two years, but really began to take off beginning in year three.

Bill Pay service is a significant portion of Internet Baking expenditures. Because of this, Mr. Jones also evaluated the cost per bill pay transaction. The number of bills paid has increased 44% per year on average. Mr. Jones noted that if these transactions were processed in the Bank's branch, it would cost \$1.46 per transaction. A transaction initiated on the Internet Banking system cost \$0.64 per transaction in 2006, which is \$0.82 less than a transaction processed by a teller. For 2006, the use of Internet Banking in lieu of a teller has saved the bank \$50,187. It was not until 2005, four years after implementation that the Bank began to realize a savings over teller generated transactions.

Limitations of the Current Internet Banking System

The Internet Banking system has not been a robust delivery channel. It has always been CASB's desire to have the Internet Banking system function as a standalone branch. As noted above, the system was limited to account inquiry, balance history, account transfer, intra-bank loan payments, and bill payment. The system did not offer the capabilities of the larger banks. The main functionality that the system was lacking was the ability to open and fund new deposit accounts.

Another limitation of the current system was the inability to take loan applications online. To workaroud this problem, the Bank contracted with a third-party known as MortgageBot to provide the Bank with an online mortgage application through the Bank's website. This product worked outside of the Internet Banking system. MortgageBot brings the customer to a secure website to complete and submit a mortgage application that can be imported directly into the loan underwriting system. This system was implemented in 2005 and has been a quite costly endeavor. Over the past two years, 26% of mortgage applications submitters were received through this online channel. The average cost per loan application has been \$355.95.

New Improvements Available in the Internet Banking Environment

After some research, Vice President of Information Technology Preston Smith presented the Information Technology Committee with a list of Internet Banking enhancements that would

help the Bank achieve its goal of making the online banking a full service banking channel. He felt the following products would be the most beneficial to CASB: 1) bill presentment and payment; 2) funds transfer; 3) check reorder; and 4) new account opening and funding. Mr. Smith provided additional detail on each product.

Bill Presentment and Payment

Mr. Smith informed the committee that research indicates that more than 50% of all households will pay their bills online by 2010. He stated that online bill payers have been shown to carry higher deposit and loan balances, are more likely to stay with their Bank, and are 287% more profitable than customers not using online banking. While the Bank already provides customers with bill payment services, bill presentment and payment will enable customers to receive and view their bills via the Internet Banking system. Not only does this service save the customer time, but the service eliminates paper bills received via the US Mail, also providing the customer with additional security from identity theft. Mr. Smith added that while the bank currently does not charge for bill payment services, it may be able to charge a nominal monthly fee to customers for bill presentment services.

Funds Transfer

Mr. Smith stated that currently consumers are able to transfer funds between their own CASB accounts and business customers are able to process wire transfer requests. He stated that an enhancement is available to enable consumers to transfer funds to individuals, small businesses, and other accounts within or outside the Bank. This service takes advantage of the Automated Clearing House Network (ACHN). The ACHN provides inter-bank transfer of funds at a significantly lower cost than a wire transfer.

Check Reorder

Mr. Smith next discussed the check reorder process. The current process requires a customer to bring a check reorder form to a branch. The branch employee must input the order onto the check vendor's order system through the company's website. Enhancements to the Internet Banking system would allow the customer to sign onto the Internet Banking system to reorder checks. The customer's order would be automatically transmitted to the check vendor. Mr. Smith stated that this would eliminate bank personnel from the check reorder process, thus freeing time for other tasks.

New Account Opening and Funding

Preston Smith stated that the current system does not provide the Bank with the ability to open accounts online. In 2005, the Bank made available an online account application to customers. This enabled the customer to complete the application online after which it would be automatically emailed securely to a customer service representative for processing. The new account representative must input the account details on the system and schedule an appointment with the customer to come to a branch, sign the papers, and fund the account. One of the key reasons a customer must come to the Bank is because the USA Patriot Act requires the customer to positively identify new customers via a driver's license and one other form of identification. This is still a manually intensive process and takes approximately 10 minutes for a customer service representative to complete. Mr. Smith stated that the online account opening and funding would automate the entire new account process. The Internet Banking system would be able to accept and approve the application, populate the core processing system with the account data and also fund the account electronically via an ACHN or credit card. The system will perform a credit check on the customer and use automated identity verification tools to verify the customer's authenticity. In accordance with the requirements of the Patriot Act, the system will

also run the customer's information against government lists to search for terrorist or other individuals or organizations banned from doing business in the United States. All activities in account opening and funding take place in a real time environment.

Security Considerations in the Internet Banking Environment

Presently the Internet Banking system requires a user ID and a password to access the system. This is known as single factor authentication. The Federal Financial Institutions Examiner's Council (FFIEC) recently issued a regulatory guidance entitled Authentication in an Internet Banking Environment. The guidance requires that the Bank perform a risk assessment of Internet Banking services and evaluate the protection of the data contained within. The guidance further states that single factor authentication generally does not provide adequate security of customer data and suggests that multi-factor authentication be implemented to further protect customer data. To access the system, multi-factor authentication requires the use of something you know (e.g. user ID and password) in combination with something you have (e.g. Smart card or random number token) or something you are (e.g. fingerprint or iris scan). The most common approach to multi-factor authentication is to have the user use a token key, digital certificate, cookies or random number token card. The regulatory guidance stipulates that a multi-factor authentication response must be in place within the next twelve months.

Following Preston Smith's presentation, the Information Technology Committee discussed the enhancements and all agreed that they would be a big step towards the transforming the Internet Banking system into a full service branch. Senior Vice President Juliana Hatfield said that she had recently spoken with a representative from Open Solutions, the provider of the Bank's core processing system. She said the representative let her know that they were now offering an Internet Banking module that could be added to the core processing system. Ms. Hatfield said this was a very attractive option because it would offer seamless integration between the core processing system and the Internet Banking system. Executive Vice President Jason Ward instructed Vice President of Information Technology Preston Smith to send out requests for proposals to the current Internet Banking provider Digital Insight and also to Open Solutions. Mr. Ward instructed that both bids must contain a quote for multi-factor authentication

THE PROPOSAL

Both Digital Insight and Open Solutions provided the Bank with proposals for three years of Internet banking service. Preston Smith, Vice President of Information Technology reviewed the proposals and prepared the comparative analysis (see Table in next page) for the IT Committee. Both proposals included quotes for the exact same services. Regarding multifactor authentication, Digital Insight uses a cookie placed on the users PC to satisfy the multifactor authentication requirements while Open Solutions uses the random number token card approach. VP of IT Preston Smith provided the committee with a couple additional facts. Open Solutions Internet Banking module was only rolled out six months ago. He let the committee know that while the product has been through beta testing, the product is still relatively untested in the market place. He also reminded the committee of difficulties that the Bank has had with new releases of Open Solutions products in the past. Mr. Smith also let the committee know that Digital Insight was recently acquired by Intuit, the makers of Quicken and QuickBooks. He said that Intuit plans to fully integrate Digital Insight's Internet Banking product with Quicken and QuickBooks.

	Current Annual	Digital Insight		Open Solutions	
		Implement	Annual	Implement	Annual
Base Internet Banking Services	\$153,096	n/a	\$133,260	\$40,000	\$195,150
New Services					
Check Reorder		\$2,625	2,256	\$2,500	\$2,700
Account Reconciliation		2,625	12,000	5,000	10,800
Funds Transfer		4,500	12,000	2,500	2,700
Funds Xfer per txn (5000 txns)		\$6/txn	30,000	\$5/txn	25,000
New Account Open/Fund		5,625	6,300	5,000	4,800
Per new acct opened (75/mo)		\$9.75/acct	5,400	Included	
Bill Payment/Presentment		13,125	30,000	15,000	36,000
Multi-Factor Authentication		25,000		28,000	
Total New Services		\$53,500	\$97,956	\$58,000	\$82,000
Grand Totals	\$153,096	\$53,500	\$231,216	\$98,000	\$277,150
Annualized costs over contract w/implementation		\$249,049		\$309,817	

OPTIONS FOR CONSIDERATION

The Information Technology Committee must determine how to best grow the Internet Banking channel into a full service branch of the Bank. The Bank understands that it must offer Internet Banking as a delivery channel of products and services. To stay competitive, the Internet Banking system must continually be upgraded to keep pace with the ever changing technological environment. The current economic environment has put a squeeze on Bank profits. The Information Technology Committee has a budget of \$200,000 per year over the next three years for Internet Banking related expenditures.

Due to the limited budget, the Information Technology Committee must decide which options will provide the best tools to transform the Internet Banking system into a more robust delivery channel for Bank products and services. Because of regulatory guidelines, the Information Technology plans to add multi-factor authentication to which ever Internet Banking product is chosen. The marketing department estimates that, if a new account feature is offered, approximately 550 new accounts can be obtained through this channel. The \$9.75 to add each account exceeds the \$5.75 cost to open an account in a branch. Marketing estimates that approximately 1,000 customers will subscribe to Internet Banking to take advantage of the Bill Pay and Presentment feature. This translates into an estimated cost of \$10 per customer using this service. Marketing also estimates that the implementation of the Funds Transfer function will result in approximately 5,000 transactions per month at a cost of \$6 per transactions.

The committee must review the options above and determine which features will give the Bank the most “bang for the buck.”

References available upon request from Haidong Liu

THE RISK PERCEPTION OF INTERNET ADDICTION:THE MODERATOR ROLE OF ONLINE SUPPORT

I-Ling Ling

Graduate Institute of Marketing and Logistics Management, National Chiayi University,
580 Sen-Min Road, Chayii City 600, Taiwan ROC

yiling@mail.ncyu.edu.tw

Chih-Hui Hsiao

Graduate Institute of Marketing and Logistics Management, National Chiayi University,
580 Sen-Min Road, Chayii City 600, Taiwan ROC

chihhui@mail.ncyu.edu.tw

Chiang Hsu

Chang Jung Christian University

396 Chang Jung Rd., Sec.1, Kway Jen, Tainan 71101, Taiwan

Chiangh@mail.ncku.edu.tw

ABSTRACT

This article studies the self-diagnosis inventory and online social support can influence the risk perception of Internet addicts. In study 1, we show that when there is a lack of contextual cues and when information about risk factors (e.g., Internet addiction) is difficult to activate from memory, people can retrieve their frequent Internet usage to assess the risk estimation and overestimate the likelihood of the risk. We label this as the *proxy diagnostic information*. In study 2, the results demonstrate that online social support moderates the effect of the self-diagnostic inventory in the self-positivity bias of Internet addiction.

Keywords: risk perception, Internet addiction, online support, self-positivity Bias

THEORETICAL FRAMEWORK

Proxy Diagnostic Information

Feldman and Lynch(1988) introduced the accessibility-diagnostic framework, which assumes an input is used in judgments if it is not only accessible but also diagnostic. Accessibility refers to the ease of retrieving the input from memory; meanwhile, diagnosticity refers to the

capability of the retrieved input to arrive at a solution for the judgment task at hand. The greater the accessibility and diagnosticity of input for a judgment relative to alternate input, the more likely it will be used [6]. Information cues that are either retrieved from memory or provided in the context lead to risk perception. When memory-based information is accessible and diagnostic, contextual information should not be used. However, when memory-based information is accessible but not diagnostic, the use of contextual information can be functional to risk perception [5].

Menon and Raghubir(2003) followed to extend the Feldman and Lynch's model, namely the mere-accessibility framework. If the behaviors are retrieved easily from memory, they demonstrated that this easy task may be perceived as diagnostic and be used as an input to risk estimation. The more easily the risk factors are recalled, the higher the risk perception is perceived. Therefore, we posit even though the real risk factors are difficult to recall from memory, if the proxy risk factors are easily recalled, people will probably use this proxy information to perform their risk estimation. However, when a self-diagnosis inventory is provided as a cue, people exploit this contextual information to form their judgment [5]. As for the health-related event, we define the proxy diagnostic information as accessible but not diagnostic memory-based information. When the true risk factor information is difficult to activate from memory, proxy diagnostic information will take over to form the risk estimation

H1: Without contextual information, the memory-based information is not accessible and memory-based proxy information is easily accessible; people exploit this information to perform their risk estimation. When the contextual information is provided, the contextual information (one kind of self-diagnosis inventory) can be used to estimate the risk of getting Internet addiction.

The Effect of Social Support on the Self-Positivity Bias

Social support refers to the resources that are provided by other persons. Cohen and Syme (1985) investigated the role of social support and demonstrated that social support protects people from suffering the pathogenic effects of stress. Young (1997) emphasized that social support can be formed on the basis of a group of people who one engaged in the regular computer-mediated communication with one another over an extended interval of time. The longer time users spend on the Internet, the more they use it for emotional support and the higher possibility they get Internet addicted behavior [7].

On the other hand, the self-positivity bias in personal risk perception is significant because it may seriously hinder the efforts to promote risk-reducing behavior. People always believe that others are at a higher risk of getting diseases than themselves. Lin, Lin and Raghubir(2003) argued that an event over which an individual does not have much control (e.g., cancer), does not implicate one's esteem. Therefore, self-positivity can be treated as a strategic

device that individuals use to maintain or enhance their self-esteem. We then propose that social support can be formed on the basis of a group of people who are engaged in the regular computer-mediated communication with one another for an extended interval period of time. And it can moderate the effects of the self-diagnosis inventory on the self-positivity bias.

H2: The self-diagnosis inventory does affect the self-positivity bias in the high online social support; but, in the low social support, the self-positivity bias is large when a self-diagnosis inventory is provided.

STUDY 1

Study 1 proposes that in estimating the likelihood of risk perception of Internet addiction, when there is a lack of a self-diagnosis inventory, the respondents rely on memory-based assessing their judgment. We believe that a self-diagnosis inventory provides a diagnostic source of information in forming risk perception of Internet addiction. Respondents that completed the self-diagnostic checklist early perceive a different level of risk of Internet addiction, compared to those who completed the checklist late (hypothesis 1).

Method

Respondents. Participants were comprised of ninety-six undergraduates in an Asian university. Respondents were assigned randomly to one of two between-subjects conditions.

Design and Procedure

A two-level between-subjects design was used to test H1. Respondents in the first condition (read DSM IV checklist first) were asked to complete the eight-item pathological checklist of Internet usage before they rated their possibility of getting Internet addicted. Respondents in the second condition(read DSM IV checklist after) were informed that the study was related to Internet addiction awareness, and they were then asked to rate their estimates of Internet addiction. After finishing their estimates, respondents completed the DSM IV checklist. The format of the DSM-IV checklist was a “Yes-No” checklist.

Dependent Measures. To assess the self-estimates of the likelihood of getting Internet addiction disorder, respondents were asked how much Internet addicted they considered themselves to be from 0% to 100%. Subsequently, respondents were given provided some background information about their activities in regard to Internet usage. As a manipulation check, the regularity was inferred from the respondents’ reports of the usage regularity along a 4-point scale ranging from one (rarely) to four (daily use).

STUDY 2

One of the implications of the self-positivity bias is that it is a route to maintain one's self-esteem[3]. In this study, we not only examine the self-positivity bias on judgment of the risk perception of getting Internet addiction, but also propose that online social support has a positive relationship with Internet addiction. We demonstrate that the self-diagnosis inventory does not affect the self-positivity bias in the high online social support; but, in the low social support the self-positivity bias is large when the self-diagnosis inventory is provided (hypothesis 2).

Method

Respondents. Ninety-five undergraduate students from an Asian university participated in the exchange for course credits. Respondents were assigned randomly to one of the two between-subjects conditions.

Research Design

We conducted a 2(self-diagnosis inventory: DSM IV present/absent) \times 2(online social support: high/low) between-subjects design. The remaining details were similar to those in study 1.

Manipulation. As in study 1, a manipulation check on Internet usage was conducted. Regularity was inferred from the respondents' reports of their usage regularity along a 7-point scale ranging from 1(very irregular) to 7(very regular). Furthermore, as for online social support measurement, respondents rated how they feel according to the social support from their Internet activities, on a 5-point semantic-differential scale (anchored at "not at all" and "very"). The scale was developed by Cohen et al. (1985) with higher scores indicating higher online social support. Factor analyses produces only one factor. Each item is summed into a single index, and the scale is reliable ($\alpha = .925$).

GENERAL DISCUSSION

Based on the results from the two studies, we find support that the self-diagnosis inventory can serve as a primarily diagnostic function for people to decide whether they are at risk. Online social support is shown to be an invisible tow to influence people to adjust their own risk perceptions as well as the risk perceptions of others. In study 1, we show that when there is a lack of contextual cues and when the information about risk factors (e.g., Internet addiction) is difficult to activate from memory, people may retrieve their frequency behavior to assess the risk estimation and overestimate the likelihood of the risk. We label this as the *proxy diagnostic information*. In study 2, we repeat the research of the self-positivity bias in the context of Internet addiction, and demonstrate that online social support moderates the effect of the self-diagnostic inventory in the self-positivity bias of Internet addiction.

Does the use of the self-diagnostic inventory lead to reasonably accurate evaluations in many health-related cases? The findings of study 1 reveal that the higher the symptoms the respondents perceive, the higher the possibility of Internet addiction they assess. Obviously, the self-diagnosis inventory dominates how people perceive their risk of Internet addiction. When the self-diagnosis inventory is a cue, the respondents use this information to assess the risk of Internet addiction. However, through their research on the persuasive effects of self-diagnosis, Raghubir and Menon (2005) showed that the manner of construction of a self-diagnosis inventory can increase the likelihood that respondents at risk identify their depressive symptoms, assess their risk and seek treatments.

Another interesting question is here: In the current context of Internet addiction, why did the respondents who had read the self-diagnosis checklist have a decreased perception of their own risk? Our position is that in general, the use of a self-diagnostic inventory should be regarded as rational. In our case, the manipulation check show that Internet usage is a regular and frequent behavior in people's daily lives. There is also a significant relationship between spending time on Internet and the risk perception of Internet addiction. As for our case, without any given contextual information, respondents can overestimate their risk of Internet addiction due to memory-based proxy diagnostic information. In contrast, if the contextual information is provided as a diagnostic cue, people can reevaluate their risk of Internet addiction. In summary, the design of a self-diagnosis inventory is able accurately identify the behavioral symptoms. Our findings are consistent with the work on the accessibility-diagnostics framework (Feldman and Lynch 1988). More importantly, without the self-diagnosis inventory as a cue, people can probably activate their memory-based proxy diagnosis information.

Our findings also suggest that online social support moderates the effect of the self-diagnosis inventory on the self-positivity bias. We demonstrate in study 2 that when people depend strongly on online social support related to the given self-diagnosis inventory, they automatically adjust the risk perception of themselves and others. On the other hand, when people's online social support is low, they only reevaluate their risk perceptions when the self-diagnosis inventory is provided and this increase the self-positivity biases. To attain the specificity, we bring theories from different studies and apply them to understanding the effect of online social support. For example, Young(1998)'s research into the assessment of pathological Internet use demonstrated that the Internet itself is not addictive, but specific applications embedded within interactive features play a significant role in the development of pathological Internet usage. People may be driven by online social support towards addictive use of the Internet. In other words, when people depend on strong to online social support, this does not diminish their self-positivity bias on Internet addiction even when using the self-diagnosis inventory as a cue. The Internet as a communication platform where social and information support are believed to be the main resources exchanged among the users, so it is a highly promoted technological tool making detection and diagnosis of addiction difficult [8].

Proper diagnosis is often too complicated since there is currently no acceptable criteria for judging Internet addiction other than those listed in the Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition (DSM IV; American Psychiatric Association, 1995).

LIMITATION AND FUTURE RESEARCH

It is not without limitations. For example, we used an indirect method (e.g., correlation analysis) to figure out the relationship between frequent Internet frequency usage and the possibility of Internet addiction. However, this assessment is somewhat unreliable.

In the future, we should come to better understand some specific research questions such as the relationships among Internet, activities (chat rooms, online games, e-mail...etc.), demographic segmentation, and users' psychological control mechanisms. Future research includes the investigation of the effect of response alternatives to discriminate heavy Internet use and addiction.

REFERENCE

- [1] Cohen, S., Mermelstein, R. Kamarch, T. and Hoberman, H. M. Measuring the functional components of social support, in *Social Support: Theory, Research and Applications*, eds. I. G. Sarason and B. R. Sarason, 1985, 73-94
- [2] Feldman, J. M. Lynch, J.G. "Self-generated validity and other effects of measurement on belief, attitude, intention and behavior," *Journal of Applied Psychology*, 1988, 73 (August), 421-435.
- [3] Lin, Y. C., Lin, C.H., and Raghurir, P. "Avoiding anxiety, being in denial or simply stroking self-esteem: why self-positivity?" *Journal of Consumer Psychology*, 2003, 13(4), 464-477.
- [4] Menon, Geeta Raghurir, P. "Ease-Of-Retrieval as an Automatic Input In Judgments: A Mere Accessibility Framework?" *Journal of Consumer Research*, 2003, 30, 230-243.
- [5] Raghurir, P. Menon, G. "Depressed or Just Blue? The Persuasive Effects of a Self-Diagnosis Inventory" *Journal of Applied Social Psychology*, 2005, 35(12), 2535-2559.
- [6] Simmons, Carolyn J., Bickart, B. A. and Lynch, J.G.Jr. "Capturing and Creating Public Opinion in Survey Research," *Journal of Consumer Research*, 1993, 316-329.
- [7] Suler, J. R. "To Get What You Need: Healthy and Pathological Internet Use," *Cyberpsychological and Behavior*, 1999, 2(5), 355-393.
- [8] Young, K. S. "What makes the Internet Addictive: Potential explanations for pathological Internet Use," Paper presented at the 105th American Psychological Association Annual Convention, Chicago, IL, USA. 1997

THE IMPACT OF “GREEN COLLAR WORKERS” ON ORGANIZATIONS¹

Diane M. Harvey, Roger Williams University, 1 Old Ferry Rd., Bristol, RI 02809
dharvey@rwu.edu, 401-254-3018

Susan M. Bosco, Roger Williams University, 1 Old Ferry Rd., Bristol, RI 02809
sbosco@rwu.edu, 401-254-3175

Gregory M. Emanuele, Roger Williams University, 1 Old Ferry Rd., Bristol, RI 02809
gemanuele279@hawks.rwu.edu, 401-254-3175

ABSTRACT

The issue of Corporate Social Responsibility (CSR) has become one of importance to all types of organizations. Being considered a socially responsible company can be advantageous both as an attraction for investors as well as potential employees. One aspect of CSR is to be considered environmentally friendly. In order to promote CSR, companies realize that they need employees who believe in sustainability and environmental causes. Thus, they seek a new type of worker classified as a “Green Collar” worker. Through this research, we will attempt to better understand this trend and determine how it is being enacted in organizations.

INTRODUCTION

Environmental and sustainability concerns have become increasingly important to consumers and workers alike. These issues have traditionally been considered an aspect of Corporate Social Responsibility (CSR) as originally defined by the Caux Roundtable [8]. One specific example of environmental activity and CSR is from Fujitsu Siemens, which plans to use its green campaign as a public relations initiative as well as a means of bringing environmental awareness into the corporate community [3]. In order to achieve desired outcomes in these areas, companies are drawing on the trend of Green Collar workers. “Green” workers are being recruited by construction companies, bicycle repair shops, and engineering firms [6]. A web site, www.greenjobs.com, based in Northern California, for companies seeking employees for “ecojobs” has registered more than 5,000 job seekers since its launch in late 2004[6]. Its advertisers include companies such as General Electric and BP Amoco.

GREEN COLLAR WORKERS

The origin of this class of workers is not precisely known. It appears thus far that it began when environmental awareness became a concern of companies such as those that manufacture computers. Through the “Green PC”, organizations became aware of the importance of being environmentally concerned at the workplace. From this awareness, the search for additional technologies and methods that reduce negative impact on the environment has grown as has the phenomenon of Green Collar workers [9]. Characteristics of these workers include: Recycles household waste, cycles to work if at all possible, belongs to organizations concerned

with environmental issues, has made energy saving alterations to their home, encourages the use of energy saving technology at work, recycles paper at work, and is interested in future energy efficient technology [9]. To summarize, a Green Collar worker is an environmentalist, someone whose functions at the workplace include protecting the environment, encouraging sustainable workplace practices, and raising environmental awareness among co-workers.

Organizations and Green Collar Workers

In today's organizations, employers are recognizing the value of Green Collar workers as they attempt to become as "green" as possible. Green Collar workers are present in many jobs such as environmental consultants, biological systems engineers, environmental lawyers, technology workers and permaculture specialists. Furthermore, green jobs are being created in the clean energy sector and in the emerging green economy. Green Collar workers may not perform environmentally related jobs, per se, but may be advocates for the environment within their organizations and are cognizant of the organization's actions with regard to environmental issues.

Green Collar workers may be more sensitive to their organization's CSR activities in the environmental arena. Research on the reactions of employees to their employer's level of Corporate Social Responsibility found that both employees' engagement levels and opinions of senior management are improved when they perceive the organization as actively involved in CSR efforts [1]. In the six countries surveyed, other positive outcomes observed were improved employee satisfaction and greater intent to remain with the organization. In addition, employees who considered their organization as being active in the area of CSR were more likely to recommend it as a good place to work. Employees also become more engaged in the organization when they perceive it as being socially responsible. They want to work for a company that is considered a responsible citizen in their community [2]. According to a 2004 Deloitte & Touche survey [11] assuming equal location, pay, benefits and responsibilities, 72% of Americans said they would prefer to work for a firm that supports charitable causes. More than half of U.S. workers want their employers to be more environmentally friendly and only 22% believe their company is doing enough or too much in this area [10]. Other surveys reveal conflicting information, however. Although more job seekers want to work for environmentally friendly companies, opportunity for advancement is their top concern with environmental policy ranking 18th of 20 items [5].

Younger employees want to work at companies that pay attention to stakeholders [4]. Organizations must be cognizant of the needs of these younger workers - the number of Gen Yers in the working world has grown 9.2% since 1999, the number of Gen X workers remained flat, and the number of Baby Boomers declined 4.3% [4]. Age related differences were also found in an international survey [1] with the oldest workers in India and China being most supportive of an employer's CSR activities while in Germany, the United Kingdom, and the U.S. it is the youngest workers.

The idea of aligning the goals of the business, employees and society is not new [7]. However, the measurement of the outcomes of these common goals in the area of environmental sustainability has not been accomplished on a widespread basis; therefore, we continue to rely on primarily anecdotal evidence about their existence and growth.

For this research we used a web-based survey to measure employee and employer trends in the areas of Green Collar workers and environmental corporate social responsibility. Employee activities measured were based on the research [9] that developed behavioral characteristics of these workers. We will also seek to compare workers in different generations to attempt to validate the findings of the earlier age-related research noted.

METHODOLOGY

Subjects were recruited using a snowball sampling method to obtain a generationally representative subject pool. This resulted in a pool of 256. There were 64% female and 36% male respondents. The generational composition was Veterans 10%, Boomers 37%, GenX 17% and GenY 36%. The percentage of workers in each generation is approximately 10% Veterans, 44% Baby Boomers, 34% GenXers, and 12% GenYers. Our sample had a more Gen Y respondents since it included students as well as employees. Employment status was as follows: Students 31%, employed in service industry 43%, employed in non-service industry 17%, unemployed 3%, and retired 6%.

The survey was designed using questions from a pilot study by one co-author as well as from surveys cited in the literature review. The survey was approved by the Human Subjects Review Board of the university. It was distributed online using Qualtrics software.

RESULTS

Knowledge of environmental issues

Most respondents, 67%, did not know what a Green Collar worker was prior to this survey. Familiarity with the Green Jobs Act of 2007 was also lacking with only 17% stating they were either very or somewhat familiar with this legislation. Regarding the use of hybrid vehicles, knowledge levels were better, with 85% indicating they felt they knew enough to state their opinion. When asked about the Kyoto Treaty, most respondents, 65%, felt they were not informed enough to answer. When asked whether methane or carbon dioxide emissions were more hazardous to the environment more responded incorrectly, with 54% choosing carbon dioxide. There were no significant differences among generations with regard to knowledge of these areas. The issue of global warming has been highly publicized, especially with the acclaim received by Al Gore's Nobel-winning film "An Inconvenient Truth". Significantly more Generation Y respondents believe that global warming is real ($F=2.70, p<.05$).

Personal Attitudes toward Environmentalism

We wanted to determine current individual attitudes about the environment to gauge the level of interest. Respondents indicated their agreement with the statement "I consider myself to be an environmentally friendly person. Most, 64% either strongly agreed or agreed. Only 5% either disagreed or strongly disagreed. We used a crosstabs to analyze generational differences, for which the chi-square was significant ($X=25.85, df=12, p<.02$). The largest segment to consider itself environmentally friendly was the Veterans of whom more than 80% agreed or strongly agreed that they are environmentally friendly. Nearly half of the younger generations were

neutral, whereas less than one-quarter of older generations chose this response. When asked whether they were willing to pay more for green technology if it benefits the environment 67.2% strongly agreed or agreed that they would do so. There was no generational difference for these responses.

To determine whether the behaviors of a Green Collar worker are being practiced, we asked respondents to indicate whether they participated in any of these activities. The majority, 91%, shut off lights and water after use, and 87% recycled. Very few, 3%, drive a hybrid vehicle. To determine whether there were any generational differences in “green collar” activity we created a variable, Green Collar Index. It was calculated by summing the number of actions performed, with higher values indicating more Green Collar characteristics. These totals ranged from 1 to 6. The crosstabs analysis by generation resulted in a significant chi-square ($X=25.66$, $df=15$, $p<.05$). The overall mean for the Green Collar Index was 2.91. The highest percent of participants was the Veterans at 56%, followed by the Boomers at 42.4%. Those with indices higher than the mean were the Boomers with 33.4%, followed by the Gen Xers with 26.1%.

Interest in Organizational Environmental Activities

To determine whether people feel it is important that the organization cares about the environment, we measured responses to the statement “It is important to me that my school/company cares about the environment”. Most, 89%, agreed or strongly agreed. There were no generational differences.

We then asked whether the respondents’ workplace or school has a department responsible for environmental/sustainability issues. Slightly more than one-third, 38%, have such a department. Nearly as many, 35%, do not. More than one-quarter, 27% did not know whether their organization has such a department.

We used a correlation to determine whether any relationship exists between an employee’s attitude toward their workplace or school and their environmental concerns. Significant correlations are indicated in the Table 1 below.

TABLE 1

Correlations between Personal and Organizational Environmental Attitudes

Variables	1	2	3	4	5
1. Age	-				
2. Work environmentally friendly	.027				
3. Recommend workplace/school	.210(**)	.381(**)			
4. Important that workplace/school cares about environment	-.056	.200(**)	.106		
5. I am environmentally friendly	.227(**)	.065	.058	.285(**)	-

** Correlation is significant at the 0.01 level (2-tailed).

DISCUSSION AND CONCLUSIONS

From these results, it is apparent that many people are still uneducated about current legislative environmental issues. The Green Jobs Act and the Kyoto Treaty were unfamiliar to the majority of participants. Knowledge about hazardous emissions was also lacking since most respondents did not know that methane emissions are more harmful than those from carbon dioxide. Information levels were better regarding hybrid vehicles with 85% of participants indicating they had sufficient knowledge to make a choice either for or against purchasing such a vehicle. Thus environmental topics that are perceived as having a personal and direct economic effect become more salient to the individual.

Current literature suggests that younger people are more interested in the environment and resource conservation. In this study, the reverse was true, more Boomers and Veterans identified themselves as being environmentally friendly. This outcome can be viewed in one way as benefiting the environmental area of CSR since these older workers are more likely to be part of organization management, and thus in a position to impact these activities. Alternatively, it could be seen as an awareness of the personally negative impact of wasting resources and money or a realization that future generations, including their children and grandchildren, may be harmed by poor environmental management.

In the workplace, there are energy saving behaviors that have been deemed characteristic of Green Collar workers. Our results indicate that these behaviors are not limited to only younger people. Organizations that act on the premise that only extremists are interested in their CSR activities with regard to environmental causes should rethink their assumption in this regard. In this study, the vast majority, 89% of all respondents, regardless of age, indicated that it is important to them that their company/school cares about the environment. The potentially helpful outcomes for organizations when they are perceived as being environmentally concerned were clearly seen through the correlation analysis. What should be of particular interest to both workplaces and institutions of higher education is the relationship between the importance of the organization caring about the environment and the willingness to recommend that organization to others as an employee/student. This relationship was significant and positive; therefore, when members are concerned about the environment, and that concern is reflected in the organization's activities, it affects the perception of the individual in a positive manner. This positive perception has the effect noted earlier of improving the individual's satisfaction with the organization as seen through his/her willingness to bring others into it.

Of course, one study cannot definitively provide evidence that organizations of all kinds should take their CSR activities with regard to the environment more seriously. There were two limitations to our study that affected its generalizability. One is that the sample was not random; therefore, the subjects were more homogenous than they would have been under random selection conditions. Second, although using a web-based instrument resulted in a subject pool distributed adequately among the generations; this method did necessarily limit respondents to those with access to technology. These results do signal to organizations that many of their members want them to be environmentally responsible and would support CSR activities in this area. There are positive outcomes to be had by organizations that act upon this support to become more visibly active when conducting environmentally friendly programs. These

outcomes include improved corporate image, reduced energy costs, and employee perception. Employee recommendations provide one of the best sources of future employees, an important component for organizational success. For institutions of higher education, student recommendations should not be discounted either. Students become alumni who not only provide a pipeline for future students, but valuable contributions to support their continued existence as well. Both would do well to pay attention to Green Collar workers – they will have a continuing impact on their organizations.

REFERENCES

- [1] Being socially responsible has a positive impact on employees as well as their local communities and the environment. (2007). *PR Newswire*. Retrieved on October 10, 2007 from <http://0-proquest.umi.com.helin.uri.edu:80/pqdweb?did=1353532291&Fmt=7&clientId=65345&RQT=309&VName=PQD>.
- [2] Engaging employees through social responsibility. (2007). *Leader to Leader*, (46), 56-58.
- [3] Fujitsu Siemens plans 'green at work' drive. (2005). *Marketing Week*, 4.
- [4] Grow, B., Hamm,S. & Lee,L. (2005). The debate over doing good. *Businessweek*, (3947), 76.
- [5] Gurchiek, K. (2007). Advancement trumps green as recruitment tool. Retrieved October 23, 2007 from http://www.shrm.org/hrnews_published/articles/CMX_023384.asp#P-8_0.
- [6] Hemlock, D. (2007). Green revolution. *Orlando Sentinel*. Retrieved October 1, 2007 from <http://innovations.harvard.edu/news/39521.html>.
- [7] Kissida, M. R. (2007). Q: How do I engage employees in the CSR agenda? *Strategic HR Review*, 6(6), 7.
- [8] Principles of Business. (1994). Caux Roundtable Principles of Business. Retrieved September 30, 2007 from <http://www.cauxroundtable.org/principles.html>.
- [9] The emergence of the green collar worker. (2007). Retrieved October 10, 2007 from <http://www.fujitsu-siemens.co.uk/rl/news/241005.html>.
- [10] Tyler, K. (2006). Going green. *HR Magazine*, 51(10). Retrieved January 2, 2007 from http://shrm.org/hrmagazine/articvler/1006/1006agenda_AI.asp.
- [11] Wilson, M. (2005). Doing good is more than a feel-good option. *Chain Store Age*, 81(10), 77.

¹ This project was supported by Roger Williams University through the 2008 Provost's Fund for Undergraduate Research

EFFECTIVENESS OF COMBINING BONUS PLAN AND BALANCED SCORECARD

Chung-Jen Fu, Department of Accounting, Phone: 886-5-5342601#5341
National Yunlin University of Science and Technology, fucj@yuntech.edu.tw

Bao-Guang Chang, Department of Accounting, Tamkang University
Phone: 886-2-26215656#2887, Email: baog@mail.tku.edu.tw

Shaw K. Chen, College of Business Administration, University of Rhode Island
Phone: 401-874-4339, Email chenshaw@uri.edu

ABSTRACT

This study uses an explicit function form to show how to construct an optimal compensation model by integrating the bonus plan and the balanced scorecard system. Model analysis shows the optimal bonus scheme for different tiers and departments with various characteristics. Meanwhile, it shows that the weight of stock granted in the compensation system would be lower when the performance evaluation system is well operated. It is particularly important to those non-listing companies or non-for-profit organizations, as the stock bonus become less necessary when managers' effort can be well reflected by the performance evaluation system. The result can be served as a reference model to design better management system.

Key words : Bonus Plan, Balanced Scorecard, Performance Evaluation, Compensation design

I. Introduction

Both the bonus scheme and the stock grants system are effective management tools to attract and retain the outstanding executives and employees. Since the advocacy of Kaplan and Norton (1992, 1993), the Balance Scorecard (BSC) has become not only a main stream of performance evaluation but also an effective tool in assisting the implementation of organization strategies. The agency theory implies that executives incentive compensation and performance measurement system are endogenously determined. If we can integrate the bonus scheme and the balanced scorecard system effectively, then we can construct a better performance evaluation and compensation system feasible for both the profit and non-profit organizations.

Previous studies of managerial compensation focus on the incentives provided to chief executive officers (CEOs). Equity-based compensation, including stock options and with/without restricted stock award, has become an increasingly important component of executive compensation (Yermack 1995; Bryan et al., 2000; Kang, et al., 2006). An important issue is how to align incentives and responsibilities of managers with the objective of firm value maximization. Aggarwal and Samwick (2003) show that pay-for-performance incentives differ across executives according to their responsibilities and are structured to motivate managers, subject to the precision with which shareholders can measure each manager's effort.

Traditionally, most companies employed financial performance measures to evaluate and reward managers (Eccles et al., 1992). Some firms would add market measures in the compensation contracts (Antle and Smith, 1986; Lambert and Larcker, 1987). Recent research shows that an increasing number of firms have incorporated nonfinancial measures such as quality, customer satisfaction, and market share into performance measurement and compensation plans (Ittner et al., 1997; Banker et al., 2000).

Firms that emphasize the short-term financial performance fail to link the long-term strategy with its short-term actions, causing a gap between strategy development and its implementation. The BSC complements the financial measures with non-financial measures, which in turn drives future financial performance. Thus, the BSC translates an organization's strategic objectives into a comprehensive set of performance measures (Kaplan and Norton, 1992, 1993). The BSC has four major performance measures: the financial perspective, the customer perspective, the internal business process perspective, and the learning and growth perspective. By measuring performance from multiple perspectives, BSC could guard against short-term oriented and sub-optimized behaviors under the traditional financial measurement system. Although a growing number of firms have adapted balanced scorecards for compensation purpose, relatively little is known about the implementation issues associated with scorecard-based reward systems (Ittner et al. 2003). Particularly, McKenzie and Shilling (1998) suggest that a difficult decision is to assign weights to performance measures in a BSC. Empirical research shows that weights on multiple performance measures for incentive compensation are often based on subjective judgment (Krishnan et al. 2005). But Ittner et al. (2003) shows that the high level of subjectivity in the BSC plan could cause the system ultimately to be abandoned.

This research extends Holmström and Milgram (1991) and Fu et al. (2002) to integrate the BSC system into linear compensation contracting model. It shows how to integrate the multiple performance measures into a representative performance index for the listed company, with a guiding principle of bonus stock (or stock option) to executives in different tiers or different departments with the company. Consistent with Aggarwal and Samwick (2003), the model analysis indicates that compensation incentives vary across executives, according to their responsibilities and the degree of precision with which shareholders can measure the effort of managers. More importantly, we show that the weight of stock bonus plan in the entire compensation system could be reduced when the effort input of executives can be well reflected from the performance evaluation system. In contrast, the weight of stock bonus plan needs to be increased if the performance evaluation system fails to reflect the effort input of executives. This dynamic relation between bonus plan and performance evaluation system provides a new venue to examine the efficiency of the design and implication of firm compensation system.

The remainder of this paper proceeds as follows. Section II develops a theoretical model to show the linear relationship between executive's compensation and multiple indicators of performance. Section III discusses the economic implications of model. We conclude in Section IV.

II. The Model

How to effectively integrate the optimal compensation contracting with the balanced scorecard in practice? According to the agency theory, shareholders design executive compensation contracts *ex ante* to minimize agency costs, after taking into account the impact of incentive compensation on managers' self-interested behaviors. The optimal contracting framework suggests that performance measurement and compensation plan are endogenously and jointly determined in a second-best incentive arrangement. The early agency literature in accounting and economics provides little guidance about how performance signals could be combined in constructing an evaluation measure (Banker and Datar, 1989). We follow the linear contracting agency model studied by Holmstrom and Milgrom (1987, 1991) and use an explicit function form to show how

to integrate the bonus scheme and the balanced scorecard system.

Maximizing the wealth of shareholders is the major goal for the company, thus we can treat the firm value or the stock price as the representative of shareholders' wealth. The firm value would be affected by (1) financial performance measures like book value and accounting earning; and (2) and non-financial performance measures such as customer satisfaction, quality, delivery, and new product developments, etc. The firm value would be higher when the operating performance is better. Therefore, we can assume the relation of stock price with these elements as follow:

$$P = c_0 + \alpha^T F + \gamma^T G + \varepsilon_p \quad (1)$$

where P represents the price of stock, F represents financial performance measures vectors of m dimension, G represents non-financial performance measures vectors of n dimension, and c_0 is intercept, reflecting the effective results of other elements. α and γ are the non-negative parameter vectors of m dimension and n dimension (represent the sensitivity of financial and nonfinancial performance) ε_p denotes random errors with $N(0, \sigma_p^2)$. Equation (1) indicates that the expectation of stock price of an individual company is determined not only by financial performance, but also non-financial performances (under BSC these non-financial performances include the customer perspective, the internal business process perspective, and the learning and growth perspective).

To successfully align the interest of management team (or employee) with the stockholders, the stockholders should design the performance evaluation and compensation system in such a way that reflects the relation denoted in Equation (1). Such a compensation system could eliminate the influence of uncertainty and lower the cost of agent. Thus, the compensation contract can be represented as the equation below:

$$B = d_0 + d_p P + \beta^T F + \delta^T G \quad (2)$$

where B represents the compensation benefit of executives, d_0 denotes fixed portion of compensation, d_p denotes a parameter or a weight placed on stock price, and β and δ are the parameter vectors of m dimension and n dimension. Equation (2) includes P as one of the determining factors of managers' compensation. As in the same spirit of Lambert and Larcker (1987), the inclusion of P can help measure the effort and performance of agents from financial and nonfinancial aspects more specifically and control the influence of non-corporation elements that are brought into F and G , the increased portion information.

Assume the agent has negative exponential utility function, and as for the agent, his effort on financial performance (f) and effort on nonfinancial performances(g), the total cost of effort is $(1/2)(f^T f + g^T g)$. The performance of financial and nonfinancial aspects of corporations (F and G) closely relates to the effort of agents (f and g). We assume the relationship among them as follow:

$$F = f + \varepsilon_F, \quad \varepsilon_F \sim N_m(\mathbf{0}_m, V) \quad (3)$$

$$G = g + \varepsilon_G, \quad \varepsilon_G \sim N_n(\mathbf{0}_n, W) \quad (4)$$

where $\mathbf{0}_m$ and $\mathbf{0}_n$ represent the null vector of m dimension and n dimension, V , W are the covariance matrix of m dimension vector ε_F and n dimension vector ε_G . Generally, although it indicates a positive relation with F and G , the relationship among measure errors are not actually related. We assume that V , W and σ_p^2 are all independent in this study.

Meanwhile, the shareholders' key issue is to determine a compensation plan $(d_0, d_p, \beta, \delta)$ that maximizes the expected residual under the incentive compatibility (IC) and individual rationality (IR) constraints. Under the linear compensation contracting, and the assumptions of manager's negative exponential utility function and normal distribution, we formulate the shareholders' issue as follows:

$$\begin{aligned} & \underset{d_0, d_p, \beta, \delta}{Max} E[P - (d_0 + d_p P + \beta^T \mathbf{F} + \delta^T \mathbf{G})] \\ s.t. \quad & E[-\exp\{-r\{(d_0 + d_p P + \beta^T \mathbf{F} + \delta^T \mathbf{G}) - \frac{\mathbf{f}^T \mathbf{f} + \mathbf{g}^T \mathbf{g}}{2}\}\}] \geq u_0 \quad (5) \\ & (\mathbf{f}, \mathbf{g}) \in \underset{\mathbf{f}, \mathbf{g}}{\operatorname{argmax}} E[-\exp\{-r\{(d_0 + d_p P + \beta^T \mathbf{F} + \delta^T \mathbf{G}) - \frac{\mathbf{f}^T \mathbf{f} + \mathbf{g}^T \mathbf{g}}{2}\}\}] \end{aligned}$$

where r ($r > 0$) represents the manager's risk aversion coefficient, \mathbf{f} and \mathbf{g} denote the feasible levels of f and g , respectively.

We assume that shareholders are risk neutral. Following the standard first order condition and backward induction method, the optimal weights assigned to the market, financial and non-financial performance measures, and the optimal agent's level of effort in financial and nonfinancial performance are obtained as follow:

$$d_p^* = \frac{C + D}{C + D + r\sigma_p^2} \quad (6)$$

$$\beta^* = \frac{\alpha}{1 + r\mathbf{V}} \frac{r\sigma_p^2}{C + D + r\sigma_p^2} \quad (7)$$

$$\delta^* = \frac{\gamma}{1 + r\mathbf{W}} \frac{r\sigma_p^2}{C + D + r\sigma_p^2} \quad (8)$$

$$f^* = \frac{\alpha}{1 + r\mathbf{V}} \left(1 + r\mathbf{V} \frac{C + D}{C + D + r\sigma_p^2}\right) \quad (9)$$

$$g^* = \frac{\gamma}{1 + r\mathbf{W}} \left(1 + r\mathbf{W} \frac{C + D}{C + D + r\sigma_p^2}\right) \quad (10)$$

where $C = \frac{\alpha^T \alpha r\mathbf{V}}{1 + r\mathbf{V}}$, $D = \frac{\gamma^T \gamma r\mathbf{W}}{1 + r\mathbf{W}}$.

From the analysis above, we have the following proposition:

Proposition:

Under the assumptions of specific utility function and cost function, the optimal weight of market performance measure (d_p^*) in compensation contract (1) is negatively associated with its variation(σ_p^2) and manager's risk aversion (r); and (2) is positively related with the sensitivity of financial and non-financial performance (α and γ); and (3) is positively associated with the measure error of manager's level of effort (\mathbf{V} and \mathbf{W}). When the financial and non-financial performance measures are informative, the optimal weight of financial (non-financial) performance measure β (δ) is (1) negatively associated with its measure error \mathbf{V} (\mathbf{W}), and (2)

positively related with the sensitivity of financial (non-financial) performance α (γ), and (3) positively associated with the measure error of manager's non-financial (financial) effort \mathbf{W} (\mathbf{V}). And both these two coefficients are positively associated with the variation (σ_p^2) of market performance measurement.

III. Economic Implications

We obtain the following corollaries related to the measurement coefficients, d_p^* , β , and δ :

Corollary 1: When an organization or a company has no market performance measure, then its $\sigma_p^2 \rightarrow \infty$. Under this condition we have $d_p^* = 0$, $\beta^* \rightarrow \frac{\alpha}{1+r\mathbf{V}} \equiv \beta'$, $\delta^* \rightarrow \frac{\gamma}{1+r\mathbf{W}} \equiv \delta'$. Then equation (2) will reduce to $B = d_0 + \beta'^T \mathbf{F} + \delta'^T \mathbf{G}$.

This indicates that setting financial and multiple non-financial performance measures as the basis of compensation is appropriate when a company has no market performance measure (such as a non-listing company or a non-profit organization) or has very high measure errors of its market performance measure to evaluate the effort that executives or departments give.

Corollary 2: For those companies the stock market performance provides a near-perfect measure, then $\sigma_p^2 \rightarrow 0$, $d_p^* \rightarrow 1$, $\beta^* \rightarrow 0$, $\delta^* \rightarrow 0$. All coefficients of financial and non-financial performance measures are approaching to zero, simple because the market price is sufficient statistic of these measures. Then equation (2) reduces to $B = d_0 + d_p P$ (indeed $d_p^* \rightarrow 1$), which is consistent with Homlström (1979) when the additional signals are not informative.

Corollary 3: For $\infty > \sigma_p^2 > 0$, $d_p^* < 1$, $\beta^* > 0$, $\delta^* > 0$, these two coefficient vectors of additional information \mathbf{F} and \mathbf{G} are positive when perfect observation of agent's action is precluded (\mathbf{V} and \mathbf{W} are positive). Both the financial and non-financial performance measures are informative. The relative weights of financial and non-financial performance measures will depend on the sensitivity (α and γ) and its measure error (\mathbf{V} and \mathbf{W}). The larger values of α to γ and \mathbf{W} to \mathbf{V} , the more sensitive B to \mathbf{F} .

For example, for senior managers with much controlling power, the measure error of financial performance factors is relatively less. Therefore, the weight of financial measure for these senior managers should be higher. On the other hand, the middle-level managers have less controlling power, so the non-financial performance measure would reflect more effort they put to the organizations. The measure error will be less than the nonfinancial measures can decide the compensation mainly.

Furthermore, under BSC structure we can decompose the nonfinancial measurement vector into customer (C), internal process (IP), and the learning/growth (LG) perspectives. The relative weights of different nonfinancial performance measures ($\delta_C, \delta_{IP}, \delta_{LG}$) are different across the departments with various controlling power and responsibilities.

Corollary 4: With the different characteristic of department, the weights of non-financial performance measures should be different. Take sales department as an example. The weights of customer perspective measure (δ_c) will be higher than internal process perspective (δ_{IP}). This is because the sensitivity of customer perspective measure (γ_c) is relative higher than internal process perspective (γ_{IP}) in sales department.

When an organization implements the balance scorecard system, it can help decrease the measure error of non-financial performance measure, in addition to increasing the non-financial performance measure systematically. As indicated from the equation (8), the optimal weight (δ) of non-financial performance measure will increase and consequently, lower the weight of market performance measure. With the continuous improvement through balance scorecard, the organization can reduce its measure error of non-financial and measure cost gradually.

Corollary 5: When a company's BSC system is functioning well, the non-financial performance measure can be a good substitute for the market performance measure; or a well-functioned BSC system can help decrease the need of stock bonus plan.

Different departments should set up their own optimal weight of performance measure. A more efficient compensation plan can help lower the loss of the departure of key employees because of the high price of stock granted. The analysis of model shows that optimal weight of performance measure can be different from the departments. The non-financial and market performance measures can be lower or even eliminated if a department or an executive can be measured through the financial performance measures particularly in short-term. The traditional compensation contract is suitable enough. On the other hand, when a department or an executive is not easy to measure with short-term financial indicators, then multiple non-financial performance measure of BSC or market performance indicator is a good candidate.

Corollary 6: The weight of market performance measure will be low and even could be eliminated if the non-financial performance measure works efficiently, as for the middle and low level of executives. The measure error is very high for their efforts to the market performance measures, thus the d_p should be lower, β and δ should be higher. In contrast, an organization could consider the bonus stock plan if the efforts of executives or employees can not be well evaluated in the short-time, such as the R&D employees or managers of a new department.

IV. Conclusions and Suggestions

Both stock granted and BSC systems have been widely used to motivate senior managers and employees. BSC can encourage executives to achieve the goal of organization by applying the multiple non-financial performance measures. Although the cost of implementing BSC is high, it benefits the firm in the long-term. The balanced scorecard, bonus and stock granted are the important tools of compensation design and incentive system. This study uses an explicit function model to show the effectiveness of integrating the balanced scorecard system with the bonus scheme to construct a better compensation structure. The results can be served as a reference model to design better management system.

References available upon request from Chung-Jen Fu

EXPLORATION AND EXPLOITATION IN THREE VALUE CREATING CONFIGURATIONS: A META ANALYSIS

I-Chung Liang, National Dong Hwa University, Taiwan, email: ichungliang@yahoo.com.tw

Yuh-Yuan Tsai, National Dong Hwa University, Taiwan, email: yytsai@mail.ndhu.edu.tw

ABSTRACT

Exploration and exploitation have emerged as the twin concepts underpinning organizational learning research. Since March (1991) explored the relationship between exploration and exploitation by means of agent-based simulation. However, the research stream is still in its preparadigmatic stage. This paper reviews past researches and discusses definitions and driving factors of exploration and exploitation as well as their relationship. In addition, it follows Stabell and Fjeldstad's (1998) typology of the value chain, the value shop, and the value network and distinguishes past researches into the three value configurations. Finally, this paper develops a number of research propositions that capture the properties of exploration and exploitation in every value creation configuration.

Key words: Exploration, Exploitation, Value Creation, Organizational Learning

INTRODUCTION

Since March (1991) explored the relationship between explorative and exploitative activities in the organizational learning process by means of agent-based simulation. Recent researches explored the same topic with different methods in different contingencies within different organizational levels or across different levels. Although they gave exploration and exploitation different definitions and explanations in different contingencies, their definitions could be covered by March's. According to March's definition, exploitation was refinement, choice, production, efficiency, selection, implementation, and execution. Its essence is the refinement and extension of existing competences, technologies, and paradigms. Its returns are positive, proximate, and predictable. On the contrast, exploration was search, variation, risk taking, experimentation, play, flexibility, discovery, and innovation. Its essence is experimentation with new alternatives. Compared to exploitation, returns from exploration are less certain, more remote in time, and often negative. Since past researches concerned different subtopics in different industries within different organizational levels, it is difficult to compare and generalize their results. This study aims to explore definitions and driving factors of exploration and exploitation as well as their relationship. In addition, it distinguishes past researches into three value creation configurations, which were proposed by Stabell and Fjeldstad (1998).

DEFINITION OF EXPLORATION AND EXPLOITATION

Exploration was characterized as search, variation, risk taking, experimentation, play, flexibility, discovery, and innovation by March (1991). Its essence is experimentation with new alternatives. Compared with exploitation, returns from exploration are systematically less certain, more remote in time, and often negative. March argued that organization absorbs the knowledge of individuals and apply them in the activities of exploration. However, rapid socialization of individuals into the

procedures and beliefs of an organization may restrain individuals from being innovative and, consequently, negatively affect the level of knowledge reflected by the organizational code. The level of organizational knowledge reflects the abundance of the sources of the organizational innovative capability. If an organization wants to increase its aggregate knowledge, it has to slow down the rate that individuals are socialized into it or speed up the rate it learns from individuals. However, since there are no obvious individual incentives for learning slowly in a group of people in which others are learning, it is difficult to slow down the rate of socialization for individuals into the organization. Moderate personnel turnover will provide opportunities for an organization to acquire new knowledge. Especially, for mutual learning has a dramatic deterioration property under conditions of exogenous environmental turbulence, moderate turnover refresh knowledge of an organization.

Exploitation was characterized as refinement, choice, production, efficiency, selection, implementation, and execution by March (1991). Its essence is the refinement and extension of existing competences, technologies, and paradigms. Its returns are positive, proximate, and predictable. March used code to represent a knowledge base of an organization and an interface for knowledge exchange between an organization and its members. He argued that individuals modify their beliefs continuously as a consequence of socialization into the organization and education into its code of beliefs, so that the knowledge of the organization can be exploited by other members. Rapid socialization of individuals into the procedures and beliefs of an organization benefits exploitation of the existing knowledge of the organization. However, it has little contribution to the level of knowledge reflected by the organizational code, since the organization has no time to absorb the knowledge of individuals. Compared to returns from exploitation, returns from exploration are systematically less certain, more remote in time. So most of organizational researchers assumed that firms tend to follow their routines, improve their existing capabilities, and squint to exploitation. In addition, there are many other factors that drive firms to exploitative activities.

RELATIONSHIP BETWEEN EXPLORATION AND EXPLOITATION

March (1991) argued that both exploration and exploitation are essential for organizations, and compete for scarce resources. Maintaining an appropriate balance between exploration and exploitation is a primary factor in organization survival and prosperity. He and Wong (2004), in their empirical test to manufacturing firms, found that the relative imbalance between explorative and exploitative innovation strategies is negatively related to sales growth rate. On the other hand, the interaction between explorative and exploitative innovation strategies is positively related to sales growth rate. According to Katila and Ahuja (2002), exploitation of existing capabilities is often needed to explore new capabilities, and exploration of new capabilities also enhances a firm's existing knowledge base. Exploration and exploitation form a dynamic path of absorptive capacity (He & Wong, 2004). In sum, exploration and exploitation are complement with each other, their interactive effect positively related to the growth of firms. However, firms may run into organizational difficulties when pursuing both strategies equally aggressively, causing the positive interaction effect to disappear (He & Wong, 2004). The causes need to be further explored.

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), Lavie and Rosenkopf (2006) found that prescriptions about whether firms should strive to manage the trade-off between exploration and exploitation are inconsistent with

observations about firms' tendencies to balance these activities in actual practice. They argued that it is because prior researches examined exploration and exploitation within a single domain, disregarding the conflicting organizational pressures that influence learning in various domains. Accordingly, Lavie and Rosenkopf examined exploration and exploitation, in the context of alliances, within function, structure, and attribute domains simultaneously. They found that firms have different tendencies within the exploration-exploitation continuum in different domains which are complement with one another. For example, a firm that shifts its focus from R&D alliances to marketing alliances over time may intensify its search for new partners and thus balance increasing tendencies to exploit in the function domain with tendencies to explore in the structure domain. Lavie and Rosenkopf also found that while explorative and exploitative activities within every domain are affected by path dependency, firms strive to adjust their policy toward the opposite direction over time and balance exploration and exploitation across domains. Their study reminds researchers to observe the tendencies of firms from different perspectives to avoid depicting only a partial picture of firms' balancing efforts.

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; Lee, Lee, & Lee, 2003; Holmqvist, 2004; Auh & Menguc, 2005; Miller, Shao, & Calatone, 2006; Perretti & Negro, 2006; Lavie & Rosenkopf, 2006). However, a group of researchers argued that some resources are unlimited, such as knowledge, information, and external resources (McNamara & Badden-Fuller, 1999; Benner & Tushman, 2003; He & Wong, 2004; Cesaroni, Minin, & Piccaluga, 2005). 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 allocated to explorative and exploitative activities according to the strategic position they chose within the continuum. In contrast, the latter believe that explorative and exploitative activities are complement with each other. Moreover, they believe that exploration and exploitation are two different and orthogonal aspects of organizational behavior and firms may acquire external resources by engaging in alliances, mergers, and acquisitions. Cesaroni, Minin, and Piccaluga (2005) observed firms' behaviors under 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 may help firms to survive. On the other had, when technological change is incremental and the firm is short of complementary assets, it can use its alliances' assets to exploit the outcome of its innovation.

VALUE CREATING MODELS VS. EXPLORATION AND EXPLOITATION

Stabell and Fjeldstad (1998) tried to explain the value chain activities of more than two dozen firms from a variety of industries by Porter's value chain framework (1985). They found that although the primary activity typology of the value chain appears well suited to describe and understand the business model of traditional manufacturing firms, it is less suitable to the analysis of activities in a number of service industries. It is not only difficult to assign and analyze activities in terms of the five generic primary value chain categories, but also often obscures rather than illuminates the essence of value crating. For example, what is raw material or product for an insurance company and a bank? What activities should the risk management of bank belong to? As the ratio of the service industry in the advanced countries is increasing, Stabell and Fjeldstad think that it is necessary to propose other value

creating models for the service industry. In following sections, characteristics of the three generic value configurations, as well as properties of exploration and exploitation are discussed.

The value chain

Porter's (1985) value chain framework is used to present and analyze the value creation at the firm level and act as a key role for the analysis of merit and flaw. In value chain models, value is created by transforming inputs into products. The value chain analysis framework postulates that competitive advantage is understood disaggregating the value creation process of the firm into discrete activities that contribute to the firm's relative cost position and create a basis for differentiation. The activities have different influence on cost and differentiation of product and on the creation of customer value. The primary activities of the value chain have both pooled and sequential interdependence and depends on one another. Therefore, they need to be coordinated. The technology development of support activities aims at reducing cost by means of process improvement or adjusting products according to the Buyer Purchasing Criteria. Similarly, every primary activity has its own explorative and exploitative activities. The effectiveness and efficiency of exploration and exploitation relies on the coordination of its adjacent activities (Siggelkow & Rivkin, 2006). Any single activity is hardly to ask its adjacent activities to cooperate with it. It has to rely on the coordination of the technology development. When technological change is fairly rapid and the dominant technology is not formed, focusing on R&D activities or looking for technological cooperation may 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 (Cesaroni, Minin, & Piccaluga, 2005).

Proposition 1: In the value chain configurations, optimal result of exploration and exploitation in any primary activity rely on the coordination of the technology development department.

Proposition 2a: In the value chain configurations, when technological change is fairly rapid and the dominant technology is not formed, focusing on R&D activities or looking for technological cooperation help firms to survive.

Proposition 2b: In the value chain configurations, when technological change is incremental and the firm is short of complementary asset, use alliances' assets to exploit the outcome of innovation make firms prosperous.

The value shop

Value shop models represent the organizational configurations which create value by an intensive technology to resolve a particular customer problem. Typical value shop firms are professional service, such as medicine, engineering, architecture, and law. Selection, combination, and order of application of resources and activities vary according to the requirements of the problem at hand. The value of the value shop is created by a variety of specialists who resolve problems together with intensive technologies. Value shop configurations can be partial activity or function within an organization. They appear as support activities in a value chain. In general, since its high coordination cost, the scale of the value shop can only create limited advantages. However, for the territoriality of clients and specialists, location can create significant advantages.

The support activities of the value shop are the same with the value chain. The tendency of firms toward exploration and exploitation is affected by selection and deployment of specialists of human resource management. The heterogeneity and foreign background of specialist within the same domain increase

the variety of resolution and the capability of innovation. A small number of experienced senior professionals lead more junior and less experienced colleagues to not only reduce the cost and increase the scale of the organization, but increase the capability of innovation for resolutions through the fresh perspectives of junior members (Perretti & Negro, 2006; March, 1991). Choice of challenging customer problems is a main means for technology development to boost the knowledge level of firms and indirectly affect capabilities of exploration and exploitation of a value shop team (Perretti & Negro, 2006). Interinstitutional problem solving teams not only can promote knowledge and problems transference between a firm and its alliances, but stimulate innovation. The interaction of members within alliance facilitates activities of exploration and exploitation. Forming an alliance with new companions benefits explorative activities. In contrast, to form another alliance with old companion or with a companion that has heterogeneous properties with other companion benefits exploitative activities (Lavie & Rosenkopf, 2003).

Proposition 3a: In the value shop configurations, the tendency of firms toward exploration and exploitation is affected by selection and deployment of specialists.

Proposition 3b: In the value shop configurations, the heterogeneity and foreign background of specialist increase the variety of resolution and the capability of innovation.

Proposition 3c: In the value shop configurations, a small number of experienced senior professionals lead more junior and less experienced colleagues increase the capability of innovation.

Proposition 4: In the value shop configurations, choice of challenging customer problems boost the knowledge level of firms and indirectly affect capabilities of exploration and exploitation.

Proposition 5: In the value shop configurations, interinstitutional problem solving team can facilitate transference and creation of knowledge within alliance and benefit exploration and exploitation.

The value network

The value network models represent firms which use a mediating technology and create value by facilitating a network relationship among customers. The typical industries of the value network configurations include transportation, insurance, and telecommunication service. In the value network configurations, firms act as mediators or agents among customers. Scale is, therefore, a potential driver of both cost and value in the value network. For example, an insurance company has to attract certain amounts of clients to share risks. However, due to the limitation of the firm infrastructure, an excessive number of clients at the same period will reduce the service quality. For example, an excessive number of connections of mobile phone in the same area will paralyze exchanges and deteriorate the quality of communication. Standardization of the format or codification of customer information will increase the efficiency of services. The business value system relationships among industry actors are not as suppliers and customers in an industry value chain, but as simultaneously coperforming levels of mediation service. For example, airports provide the infrastructure for airline companies which in turn serve as the traveling infrastructure for traveling agencies. Common industrial standard among networks will facilitate the development and utilization of demand-side economies of scale. For example, ATM services across banks facilitate the development and utilization of demand-side economies of scale.

The primary activities in the value network include network promotion and contract management, service provisioning, and network infrastructure operation. Because of the need for synchronization and dimensioning of simultaneous activities, there is reciprocal interdependence across primary activities. The support activities of the value network are the same with the value chain. Development of the network infrastructure and the new service creation from technology development increases the competitive advantage for firms. Value network configurations firms can offer new services or attract new customers on the base of the existing number and characteristics of the customer by means of introducing new technologies. Value of the value network was affected by network externality. If firms want to substitute the existing technology with the new one, the new technology has to possess higher potential for improving the performance of new products than the existing one and attract a certain amount of early birds (Lee, Lee, & Lee, 2003). Selecting new companions in the alliance benefits both exploration of new services and exploitation of existing services (Lavie & Rosenkopf, 2006).

Proposition 6: In the value network configurations, a superior new technology increases the competitive advantage of firms. However, the success of the new technology is affected by externality of network.

Proposition 7: In the value network configurations, selection of new companion contribute to exploration of new services and exploitation of exist services.

CONCLUSIONS

By means of reviewing past researches, this paper finds that March's (1991) definitions for exploration and exploitation cover definitions of other researches which explored the relationship between exploration and exploitation proposed under specific industries and contingencies. During the process of analysis, the research induces driving factors of exploration and exploitation as well as their relationship. According to the characteristics of samples and configurations of industry which the document surveyed or explored, this paper distinguishes past researches into the three value configurations proposed by Stabell and Fjeldstad's (1998). Consequently, the research proposed seven propositions to address the properties of every configuration. In the value chain configurations, optimal results of exploration and exploitation in any primary activity rely on the coordination of the technology development department. When technological change is fairly rapid and the dominant technology is not formed, focusing on R&D activities or looking for technological cooperation helps firms to survive. In contrast, when technological change is incremental and the firm is short of complementary asset, use alliances' assets to exploit the outcome of innovation make firms prosperous.

In the value shop configurations, the tendency of firms toward exploration and exploitation is affected by selection and deployment of specialists. The heterogeneity and foreign background of specialist will increase the variety of resolution and the capability of innovation. A small number of experienced senior professionals lead more junior and less experienced colleagues increase the capability of innovation. Choice of challenging customer problems boosts the knowledge level of firms and indirectly affects capabilities of exploration and exploitation. Interinstitutional problem solving team can facilitate transference and creation of knowledge within alliances and benefit exploration and exploitation. In the value network configurations, a superior new technology increases the competitive advantage of firms. However, the success of the new technology is affected by externality of network. Selection of new companion contribute to exploration of new services and exploitation of exist services.

THE INFLUENCE AND DETERMINANTS OF DIRECTORS' REPUTATION

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
Chuan-Ying Hsu, Da-Yeh University, 112 Shanjiao Rd., Dacun, Changhua 51591, Taiwan,
Tel: +886-4-851-1888#3017, cyhsu@mail.dyu.edu.tw

ABSTRACT

This paper examines the role of listed-company independent directors' reputation from the standpoints of receiving firm. Our evidence shows that director reputation is significantly related to improving firm performance in Taiwan. In addition, it can offset the negative busyness effect. In this paper, we explore the factors affecting busy directors' reputation and find that busy directors with a listed-company CEO or a college professor experience are more likely to have a good reputation, and this reputation enhances firm performance. Meanwhile high-performing receiving firms attract reputable directors to join and exert their monitoring talent to improve the quality of firms.

Keywords: Director reputation; reputation effect; busyness effect; multiple directorships

INTRODUCTION

In recent years, large business scandals have been associated with corporate governance, thus calling worldwide attention to the issue of board independence. One common criticism is that corporate boards have too few independent directors, so the practice of appointing independent directors has become widespread. Appointing more independent directors to boards to improve corporate governance in emerging markets is required. For example, Asian countries such as China, Korea, Singapore, Hong-Kong and Taiwan require a minimum benchmark on the number (proportion) of independent directors for all or newly listed companies. Therefore, the issue of demand and supply of qualified independent directors is of great concern.

Although some literature in the U.S. and Europe investigates the impact of directors with multiple directorships on performance and find the mixed results (Beasley, 1996; Loderer & Peyer, 2002; Ferris, Jagannathan and Pritchard, 2003; Fich and Shivdasani, 2006), few empirical studies have appeared in Asia. Taking advantage of Taiwan's institutional transitions, we concentrate on the presence of a busy director due to the insufficient supply of independent directors and its impact on firm performance.

In our study, a busy director is an independent director holding three or more board seats at the same time. "Busyness" is driven either by the development of the director's reputation or the perquisite consumption. Taking multiple directorships as proxies for reputational capital, some empirical studies propose that the presence of multiple directorships was positively related to firm performance or shareholder interests (Brown & Maloney, 1999; Miwa & Ramseyer, 2000). This is what we called "reputation effect", the higher the reputation, and the better the performance. On the other hand, some empirical studies show that busy directors/boards may

limit the monitoring effectiveness of outside directors and weaken corporate performance. When an independent director (or a majority of independent directors) has three or more seats, he/she is overstretched and cannot improve board quality and firm performance. This is what we called “busyness effect”.

Although quite a few articles suggest that outside directors help improve firm performance and board effectiveness (Rosenstein & Wyatt, 1990; Westphal & Zajac, 1995; Mayers, Shivdasani & Smith, 1997; Rodriguez & Anson, 2001 and Hermalin & Weisbach, 2003), do busy outside directors? We wonder whether a busy independent director does not monitor management and thereby impairs firm performance. Does a director’s reputation matter? If so, what factors influence a director’s reputation? This paper analyzes these questions and one of its main contributions is to investigate the impact of busy directors with good reputations on receiver firms appointing independent directors in emerging economies like Taiwan.

DATA AND METHODS

Sample selection and description

Because the disclosure requirement of independent directors in Taiwan starts with the 2001 annual report, our data consists of listed firms that announced the appointment of independent directors from 2002 through 2004 with available data from the Taiwan Economic Journal (TEJ) database and market observation post system provided by the Taiwan Stock Exchange Inc. All financial companies are excluded from the sample subject to the regulatory effects. This generates a final sample of 278 firms, from which we obtain directors’ names, principal professional affiliation, number of directorships, equity ownership, and so on. Table 1 presents the distribution of individual independent director announcement by year and by number of directorships.

TABLE 1

The number of directorships distribution

Panel A: Year Distribution and Frequency distribution for independent director appointment					
Directorships Held	2002	2003	2004	Number of directors	Percentage frequency
1	103	89	81	273	54.82%
2	29	33	51	113	22.69%
3	11	11	15	37	7.43%
4	7	14	10	31	6.23%
5	2	9	3	14	2.81%
6	0	3	7	10	2.01%
7	1	2	2	5	1.00%
8 and up	3	8	4	15	3.01%
Total directors	156	169	173	498	100.00%
Total directorships	284	395	383	1,062	
Number of firms	91	94	93	278	

Methodology

We test the effect of busy director/board and director reputation on firm's performance and identify the determinants of reputation by using the following regressions:

$$Y(\text{Performance}) = \beta_0 + \beta_1 \text{BusyDir/BusyBrd/MultiPCT} + \beta_2 \text{ReputCap} + \beta_3 \text{IndepPCT} + \beta_4 \text{DirOwnsp} + \beta_5 \text{LnBoard} + \beta_6 \text{LnAssets} + \beta_7 \text{Growth} + \beta_8 \text{ROS} + \varepsilon \quad (1)$$

$$\text{Logit}(\text{ReputCap}) = \beta_0 + \beta_1 \text{ROA}_{t-1} / \text{MB}_{t-1} + \beta_2 \text{LnAssets} + \beta_3 \text{LnBoard} + \beta_4 \text{IndepPCT} + \beta_5 \text{Growth} + \beta_6 \text{Public} + \beta_7 \text{CEO} + \beta_8 \text{Professor} + \beta_9 \text{CPALAW} + \beta_{10} \text{BANKINV} + \varepsilon \quad (2)$$

All related variables in both equation (1) and (2) are defined below Table 2 and Table 3, respectively.

EMPIRICAL RESULTS

Busyness effect, reputation effect and interaction effect

Model 1 of Table 2 reports that the busy director is negatively related to accounting-based ROA. In Models 2 and 3, we use different indicators to exhibit that the busier board also can destroy firm performance as posited by Fich and Shivdasani (2006). However, these three models show that directors' reputational capital is positively associated with industry-adjusted ROA. Models 4 and 5 include an interaction term between the reputation and the busy director/board indicator variable. The results confirm that the busy directors/board did harm the firm's performance, but the reputation plays the key role to alter this situation. In other words, busy directors/boards with good reputation can enhance firm performance.

As a further test of these effects, we substitute market-based variable (not shown), the industry-adjusted market-to-book value, for industry-adjusted ROA. Still there is a negative relationship between busy director/board and firm value, but a significant positive relation exists between reputation and firm value.

The determinants of reputation

From the receiving firm's perspective in Tables 2, we find that reputation signals the quality of a director and its influence outweighs the negative busyness effect. This study thus explores what factors affect reputation. In Table 3, Models 1 & 2 show that previous accounting performance in receiving firms may slightly strengthen a quality director's desire to join their boards and meanwhile CPAs or lawyers experience gain more chances to obtain external board seats. In addition, CEOs in non-public corporations or non-CEOs in publicly-traded companies might not add enough value for firms.

By replacing accounting-based performance with market-based benchmark in Model 3, we find the larger the market-to-book value of a receiving firm, the better the possibility of a good reputation. This finding implies that high-performing receiving firms draw attention of reputable directors and thus present a positive relation between these two elements. An alternative interpretation is that directors with good reputations are less likely to accept directorships in

TABLE 2

Impact of directors' reputation on accounting-based firm performance

	Dependent Variable: Adjusted-ROA				
	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	0.0609 (0.3533)	0.0821 (0.2100)	0.0924 (0.1516)	0.0669 (0.3051)	0.0902 (0.1589)
BusyDir(0,1)	-0.0434*** (0.0027)			-0.0643*** (0.0006)	
MultiPCT		-0.0495** (0.0295)			
BusyBrd(0,1)			-0.0330** (0.0240)		-0.0591*** (0.0028)
ReputCap	0.0100*** (0.0089)	0.0100** (0.0259)	0.0085** (0.0306)		
BusyDir(0,1) x ReputCap				0.0120*** (0.0029)	
BusyBrd(0,1) x ReputCap					0.0121*** (0.0043)
IndepPCT	0.0159 (0.7359)	-0.0093 (0.8404)	-0.0081 (0.8611)	0.0231 (0.6233)	0.0033 (0.9432)
DirOwnsp	0.0010*** (0.0009)	0.0010*** (0.0009)	0.0010*** (0.0007)	0.0010*** (0.0008)	0.0011*** (0.0005)
LnBoard	0.0209*** (0.0056)	0.0178** (0.0167)	0.0177** (0.0172)	0.0206*** (0.0062)	0.0176** (0.0170)
LnAssets	-0.0095** (0.0167)	-0.0103*** (0.0099)	-0.0109*** (0.0058)	-0.0091** (0.0212)	-0.0102*** (0.0091)
Growth	-0.1165** (0.0352)	-0.1197** (0.0320)	-0.1254** (0.0250)	-0.1185** (0.0316)	-0.1271** (0.0221)
ROS	0.4151*** (0.0000)	0.4133*** (0.0000)	0.4147*** (0.0000)	0.4125*** (0.0000)	0.4118*** (0.0000)
Adjusted R Square	0.5369	0.5293	0.5299	0.5404	0.5360

The dependent variable is the receiving firm's industry-adjusted ROA. BusyDir (BusyBrd) equals 1 if the board in a receiving firm has a busy independent director (50% or more of the independent directors) holding three or more seats, and 0 otherwise. MultiPCT measures the busy level for a board and is defined as the percentage of busy independent directors in corporate board. ReputCap is the director's reputational capital, measured as the average number of directorships held by an independent director. Percentage of independent directors (IndepPCT) is the number of independent directors divided by total number of board directors. Directors' ownerships (DirOwnsp) is directors' share divided by total outstanding share. Board size (LnBoard) is natural log of the number of board directors, and firm size (LnAssets) is measured by total assets. Growth opportunity (Growth) is total depreciation divided by net sales. Return on sales (ROS) is measured as operating profit divided by net sales. Statistical significance at the 1, 5 and 10 percent levels is indicated by ***, **, and *, respectively.

TABLE 3

Determinants of busy directors' reputation

	Dependent Variable: ReputCap(0,1)				
	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	4.1933 (0.3240)	4.9889 (0.2412)	-2.9342 (0.6597)	-2.6010 (0.7055)	-2.3932 (0.7229)
Industry-adjusted ROA _{t-1}	2.7533 (0.1934)	2.9922 (0.1631)			3.6451 (0.2680)
Industry-adjusted M/B _{t-1} value			2.1809 (0.0083)	2.3075 (0.0061)	2.3132 (0.0068)
LnAssets	0.2745 (0.2277)	0.2232 (0.3328)	1.1190 (0.0089)	1.1456 (0.0099)	1.2260 (0.0100)
LnBoard	-3.3445 (0.0239)	-3.2281 (0.0345)	-4.9103 (0.0788)	-4.7992 (0.0890)	-5.2861 (0.0696)
IndepPCT	-8.4514 (0.0319)	-8.6359 (0.0292)	-19.1802 (0.0322)	-19.0259 (0.0374)	-20.6294 (0.0240)
Growth	-0.6438 (0.7798)	-0.8349 (0.7186)	-23.0637 (0.0210)	-22.3649 (0.0187)	-23.8127 (0.0161)
Public(0,1)	0.9468 (0.1543)		1.7413 (0.1007)		
CEO(0,1)	0.4243 (0.5651)		1.5988 (0.3411)		
PUBCEO		1.5831 (0.0205)		2.7516 (0.0126)	2.8291 (0.0115)
Professor	0.9363 (0.2489)	0.8378 (0.2300)	2.7765 (0.1158)	1.8640 (0.0996)	2.0437 (0.0843)
CPALAW	2.5100 (0.0628)	2.3678 (0.0612)	2.7838 (0.2034)	1.7554 (0.2416)	1.5675 (0.3118)
BANKINV	1.4422 (0.1114)	1.6116 (0.0782)	0.5325 (0.6420)	0.8323 (0.4841)	0.6670 (0.5872)
Nagelkerke R Square	0.3155	0.3476	0.6235	0.6478	0.6611

The dependent variable takes a value of 1 if the number of average directorships per independent director is 3 or more, 0 otherwise. Individual characteristics adopt dummy variables defined as: Public = binary variable taking a value of 1 if the busy director serves in a publicly-traded company, 0 otherwise; CEO = binary variable taking a value of 1 if the busy director serves as a CEO and/or chairman in either publicly or non-publicly-traded company, 0 otherwise; Professor = binary variable taking a value of 1 if the busy director serves as a professor in a college, 0 otherwise; BANKINV = binary variable taking a value of 1 if the busy director is an executive in commercial banking and investment industry or works as an investment advisor, 0 otherwise; CPALAW = binary variable taking a value of 1 if the director's professional affiliation is CPA or lawyer, 0 otherwise, BANKINV = binary variable taking a value of 1 if the busy director is an executive in commercial banking and investment industry or works as an investment advisor, 0 otherwise. P-values are reported in parentheses below each coefficient estimate.

underperforming firms. CEOs serving in publicly-traded firms (PUBCEO) maintain their good reputation (Model 4), consistent with Fich's (2005) conclusion that CEOs are sought as outside directors to enhance firm value. It is more interesting to note that when we switch the performance benchmark in Model 4, the probability of obtaining good reputation is higher if the director is a professor. In Model 5, we put two performance measures together and find results similar to those in Model 4.

CONCLUSIONS

In this paper we examine the proposition that some busy directors have better reputations than others and that this can benefit the receiving firm. We selected 278 announcements from 2002-2004 and found a significantly negative relation between busy directors and firm performance. This finding is robust for different indicators for busy directors. At the same time, our results indicate significantly positive relation between reputation and firm performance. This finding is also robust for both accounting and marketing measures of firm performance.

During the period of institutional transition for independent director reform in emerging markets like Taiwan, government and public corporations perceive the importance of independent directors. However, the dearth of independent directors is a problem, so they hesitate to enforce the minimum number of independent directors. So far, only the newly listed corporations and the financial corporations are required to meet the minimum two independent directors. Naturally either the multiple directorships or the relaxing qualification for independent directors might become a common situation in Taiwan. It's a dilemma if the accumulation of board seats is harmful. Our results provide evidence that director reputation plays a role in solving this problem and explore the factors affecting director reputation. We believe that reputable directors can use their professional ability to improve corporate governance.

Furthermore, we employ the quantile function of conditional distribution to justify the robustness of our results (not shown). The breakdowns of entire distribution show that the higher reputation for busy director/board brings significantly better performance in most situations. However, the quantile regression gives more information on tail behavior of a distribution indicating that too many directorships indeed reduce the actual input and contribution of a reputable but busy director. Future research can determine the appropriate number of directorships for independent directors.

Our empirical results imply that even though the presence of multiple directorships in emerging markets may hurt firm value, the reputation is the determinant. Policy regulators should consider multiple directorships as good incentive for labor market and determine the appropriate number of directorships for independent directors. Receiving firms facing the upcoming mandate practice of appointing more independent directors in the board can refer to the factors which we provide in this study to increase the possibility of high reputation capital and thus enhance firm performance.

References available upon request from Chuan-Ying Hsu, cyhsu@mail.dyu.edu.tw.

Teaching Accounting Ethics in the Light of Bloom's
Affective Domain

Many hierarchal systems of learning have been developed over the years. Probably the most popular and oft quoted taxonomy of learning objectives is that developed by Prof. Benjamin S. Bloom of Chicago University (Kinnes, 1998-2007). Kinnes stated that the Bloom system “is easily understood and widely applied.”

Bloom, and his colleagues, actually developed three taxonomies of learning. The taxonomies were developed for the cognitive, affective, and psychomotor (skills) domains. Many educators can readily quote and apply the cognitive domain. In this article I will argue that the teaching of ethics falls within the affective domain. Therefore, the teaching of ethics can be more effective in taught using the principles of Bloom’s taxonomy for the affective domain.

What is the Affective Domain?

Bloom, etal (Krathwohl, Bloom, & Masia, 1964) described the affective domain as “objectives which emphasize a feeling tone, an emotion, or a degree of acceptance or rejection.” (p.8). The affective domain, therefore, deals with attitudes and values. These are concepts that are difficult for educators to measure and asses in their students. They are difficult to measure because they are internalized by the student.

Bloom’s Affective Domain Taxonomy has 5 major levels and then sub-levels. Unfortunately, Bloom, et al, did not give in depth explanation of each level. They are as follows:

- I. Receiving: Being aware of or attending to something in the environment
 - a. Awareness: The individual’s attention is attracted to the stimuli.
 - b. Willingness to receive: The individual is differentiates the stimuli form others and is willing to give attention.

- c. Selected attention: The individual looks for the stimuli
- II. Responding: Showing some new behaviors as a result of experience.
 - a. Acquiescence: The individual complies with expectations.
 - b. Willingness: The individual responds increasingly to an inner compulsion.
 - c. Satisfaction: The individual responds emotionally (for personal pleasure).
- III. Valuing: Showing some definite involvement or commitment
 - a. Acceptance: Continuing desire to be involved
 - b. Preference: The individual seeks out the stimuli for enjoyment
 - c. Commitment/conviction
- IV. Organization: Integrating a new value into one's general set of values, giving it some ranking among one's general priorities.
 - a. Conceptualization: Desires to evaluate the stimuli.
 - b. Organization of a value system: Gives the stimuli a value and place in his/her life.
- V. Characterization by a value or value complex: Acting consistently with the new value; person is known by the value.
 - a. Generalized set:
 - b. Characterization: Develops a consistent philosophy in life

(Definitions of major categories taken from teaching slides by William. Huitt (1998))

What is Ethics?

Ethics means different things to different people. Ethics has been described as morality, rules of behavior, or just theories (What Is Ethics All About?, 2007).

Velasquez, et al (Velasquez, Andre, Shanks, S.J., & Meyer, 1987) said that the ‘meaning of ethics is hard to pin down.’ They described ethics as “well based standards of right and wrong” and the “study and development of one’s ethics standards.”

I believe that one of the best definitions comes from the dictionary (8 Entries Found for Ethics, 2006): a study of the nature of morals and moral choices and rules or standards governing conduct or members of a profession.

In the teaching and practice of accounting both definitions should apply. It is imperative that every accountant know, understand, and apply the various codes of conduct. Codes of conduct are promulgated by the various states boards of accountancy, regulatory agencies (such as the Internal Revenue Service and the Securities and Exchange Commission), and the many professional associations. It is also imperative that accountants behave in an ethical manner. While the codes of conduct are being taught in accounting and auditing classes, ethical behavior is another issue. It is my belief that this must be taught according to teaching strategies in the affective domain.

How Ethics are Currently Being Taught to Accounting Students

Generally the ethics of the accounting profession is taught in the auditing course. While this course is required for CPA candidates, non-CPA qualifying programs may or may not have an auditing course. Therefore there are accounting students that may never get ‘ethics education.’ Furthermore, when ethics is taught in an auditing course, all too frequently it is not much more than teaching the Code of Ethics of The American

Institute of Certified Public Accountants (AICPA) and maybe some of the rules of practice of the SEC.

Lampe (1997) opined that students have to be affected emotionally as well as intellectually. However, he later talks about students having “appropriate knowledge and tools.” This very much sounds like teaching in the cognitive domain.

D’Aquila (1999) described the teaching of ethics to accounting students. She talked about laying out the ethics to be taught in the syllabus, teaching the ‘Code of Conduct,’ and using end-of-chapter questions as well as using video tapes and current events. The mainstay of D’Aquila’s article is that we should teach the ‘Code of Conduct.’

Of course it is very important for students to know this code. However knowing a list of rules is only part of what’s needed.

Why Teaching of Ethics Falls within the Affective Domain

Many of the articles written about teaching of ethics talk about something other than cognitive learning but are not specific. Lampe (1997) talked about influencing students and “values education.”

Paine, as quoted by Boatright (1991), said that the objective of ethics education is character development.

There are many articles that talk about the use of case studies as well as role-plays. The purpose of this type of education is to change student’s attitude and behaviors.

Applying the Affective Domain to Accounting Students

If the purpose of ethics education is to change attitudes and behaviors then Bloom's Affective Domain is the lens through which to develop teaching methods for the teaching of ethics. Students should be made aware of ethical issues, become involved in ethical thinking, show ethical behavior, and develop a set of ethical values and have an ethical philosophy in life.

Conclusion

Since Bloom published his book on the taxonomy of education in cognitive domain teachers at all levels have been using the taxonomy as a guide to develop lessons and set goals for their students. Bloom's taxonomy of education in the affective domain can similarly be used as a guideline for the teaching of ethics. Using the taxonomy will help to develop educational goals and hence teaching methods will be effective in reaching those goals.

Reference List

1. *8 entries found for ethics.* Web site: URL dictionary.reference.com/browse/ethics
2. *What is ethics all about?* Web site: URL www.ethics.org.au/about-ethics/what-is-ethics/what-is-ethics.html
3. Boatright, J. R. (1991). From theory to practice: faculty training in business ethics. *Moral Education Forum*, 16(4), 7-13.
4. D'Aquila, J. M. (1999). Teaching business ethics in accounting classes. *Business Education Forum*, 53(3), 20-21.
5. Kinnes, T. (1998-2007). *Benjamin Bloom and the taxonomy of learning.* Web site:
6. Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of Educational Objectives: The classification of educational goals: Handbook II: Affective domain.* New York: David McKay Company, Inc.
7. Lampe, M. (1997). Increasing effectiveness in teaching ethics to undergraduate business ethics. *Teaching Business Ethics*, 1, 3-19.
8. Velasquez, M. , Andre, C., Shanks, T., S.J., & Meyer, M. J. (1987). What is Ethics? *Issues in Ethics*, 1(1).

Despite Initial Fears to the Contrary, It Appears that Sarbanes-Oxley Gave Private Litigants a “Dull Sword” When it Comes to Piercing the Corporate Veil

Dr. Nathan S. Slavin, Hofstra University, Department of Accounting, Taxation & Legal Studies,
(516) 463-5690, actnzs@hofstra.edu

Glen M. Vogel, Esq., Hofstra University, Department of Accounting, Taxation & Legal Studies,
(516) 463-6191, lawgmv@hofstra.edu

ABSTRACT

Six years ago, while Enron crumbled and investigators uncovered Arthur Anderson’s role, the American public screamed for the government to intervene and prevent such economic disasters from happening again. On July 30, 2002, President Bush signed the Sarbanes-Oxley Act.[1] One of the major goals of Sarbanes-Oxley (“SOX”) was to create an environment of greater corporate integrity and investor confidence by holding company officers personally accountable for financial misdeeds. Sections 302 and 906 of SOX, the so-called “certification provisions” were added with the goal of subjecting corporate officers to personal civil and criminal liability unlike ever before.

One could easily believe that, in light of the above, corporate executives would be less apt to get away with certifying false or fraudulent financial information to the SEC and/or public. Despite that presumption, the question remains – is submitting a false certification under section 10(b) of SOX and SEC Rule 10b-5 enough to prove fraudulent intent and enable a defrauded shareholder to pierce the corporate veil? The question appears to be a difficult one to answer with any degree of certainty.

INTRODUCTION

Over the past seven years, accountants have managed to ride-out the storm of massive corporate accounting scandals, the demise of one of the premier accounting firms in the country, a tarnished perception of the profession, and the resulting impact of the landmark passage of Sarbanes-Oxley.[2] Even though the worst seems to be behind us, the question still remains – are there still more financial ticking time-bombs sitting out there waiting to go off and, if so, are investors better protected from

economic ruin by the likes of Sarbanes-Oxley and related regulation? Despite early predictions and hopes that this legislation would stop future scandals and restore investor confidence - the jury is still out on whether this landmark litigation truly affords investors the protections it was intended to provide.

One of the major goals of Sarbanes-Oxley (“SOX”) was to create an environment of greater corporate integrity and investor confidence by holding company officers personally accountable for financial misdeeds. The impetus behind the inclusion of sections 302 and 906 – the SOX provisions that require CFOs and CEOs to certify financial statements to be filed with the SEC – was that Congress intended these to be critical pieces of the statutory puzzle to preventing corporate corruption and fraud. These provisions subjected corporate officers to personal civil and criminal liability unlike ever before. In fact, early predictions were that these “veil piercing” provisions were going to be an effective tool for preventing future scandals and for restoring investor confidence. In reality, it appears that SOX, particularly section 302, has not made prosecuting corporate officers any easier than it was under traditional veil piercing methods. As noted by one observer, “[p]erhaps forcing top executives to certify regulatory filings is not quite the fix lawmakers might have expected.”[3]

A Look at the Pleading Requirements or “Hurdles” Faced by Litigants Attempting to Pierce the Corporate Veil

Historically, limited liability, in the corporate sense, meant that creditors of failed corporations remained unpaid. The common response in these instances was for the creditor to try to bring suit against one or more of the corporation’s shareholders and argue that liability should be extended beyond the confines of the corporate structure directly to the shareholder(s).[4] This is commonly referred to as “piercing the corporate veil.” Successfully piercing the corporate veil is often referred to as a “Herculean task” because courts are not quick to ignore the corporate formalities and extend liability directly to shareholders, recognizing that “veil piercing is an extraordinary procedure that should not be used lightly.”[5]

Actions seeking to pierce the corporate veil alleging securities fraud are typically brought under Section 10(b) of the Securities Exchange Act and Securities Exchange Commission (“SEC”) Rule 10b-5. Under these, and related regulation, a plaintiff is required to “state with particularity, facts giving rise to a strong inference that the defendant acted with the required state of mind” necessary to prove fraud.[6] “State of mind”, or scienter, is a “mental state embracing intent to deceive, manipulate or defraud.”[7] That is, a plaintiff must plead a “strong inference that the defendant[s] acted at least recklessly” in making false statements, or with knowledge of their falsity.[8] One of the main problems with the aforementioned is that nowhere in SOX does Congress define the terms “scienter” or “strong inference”. It appears that these terms were deliberately left undefined because Congress stated it did not intend to codify how these standards should be met.[9]

Prior to SOX, the U.S. Circuit Court of Appeals, for the Second Circuit, held that pleading a “strong inference” of scienter could be established by showing that the defendant either acted recklessly or had the motive and opportunity to commit fraud, or where it was alleged that corporate insiders misrepresented to the public material facts about the corporation’s performance or future prospects in order to keep the stock price high.[10] While this definition aided litigants in the Second Circuit, it was not the interpretation used by all courts. In fact, the U.S. Circuit Courts have split on these definitions with the Ninth Circuit raising the scienter element to require “strong evidence of deliberately reckless or conscious misconduct.”[11] At the same time, the Sixth and Eleventh Circuits have taken the middle road between the Second and Ninth Circuits and held that a showing of motive and opportunity alone were not sufficient to meet the scienter requirement.[12]

Recently, on June 21, 2007, the U.S. Supreme Court stepped in and addressed how to treat an examination of the pleading requirements in a § 10(b) and Rule 10b-5 action and specifically how to define a “strong inference” of scienter.[13] In determining the existence of a strong inference of scienter, courts must now consider the complaint in its entirety, as well as documents incorporated into

the complaint by reference (such as SOX certifications), and determine if an inference of scienter has as much plausibility as inferences to the contrary.

Have the Pleading Requirements for Fraud Been Impacted By the Passage of Sarbanes-Oxley?

Sections 302 and 906 of SOX were intended to be a codification of the “piercing the corporate veil” doctrine by requiring a corporate officer to expose himself to civil & criminal liability by certifying that he was directly responsible for the content of the SEC filings. These provisions were intended to make it easier for litigants and prosecutors to get to the individual officers responsible for the fraud. In fact, as noted by one recent decision; signing of the Sarbanes-Oxley certifications gives rise to an inference of scienter because they provide evidence regarding the officer’s state of mind and that the corporate officer knew or should have known about the improper financial reporting.[14]

One could easily believe that, in light of the above, corporate executives would be less apt to get away with certifying false or fraudulent financial information to the SEC and/or public. Despite that presumption, the question remains – is submitting a false certification under section 10(b) of SOX and SEC Rule 10b-5 enough to prove scienter and enable a litigant to pierce the corporate veil? The question appears to be a difficult one to answer with any degree of certainty.

For example, in August of 2007, the United States District Court for the Northern District of Ohio, Western Division, held that the signing of SOX certification reports - by the CEO and board chairman - that contained financial mistakes was not enough to sufficiently plead scienter without a showing that reports were intentionally misleading.[15] In California, the Northern District Court went further and held that a SOX certification containing false declarations is not independently actionable under section 10(b) or Rule 10b-5 but may be used merely as an “inference of scienter.”[16] More disturbing, the District Court for the District of Colorado recently went even further and not only didn’t reach the scienter question, but rather, declared that sections 302 and 906 of SOX do not even grant

parties a private right of action against corporate officers.[17] It has been noted by one Louisiana District Court that there has been really only one case in which a District Court found that a certification executed by a corporate officer pursuant to section 302 of SOX supports an inference of scienter against the certifying defendant.[18]

Perhaps forcing top executives to certify regulatory filings pursuant to SOX is not quite the solution to the corporate financial scandal problem that Congress or investors had hoped for. The certification requirement is a powerful incentive for corporate executives to do the right thing; however, because of the requirement that a plaintiff must still establish a strong inference of scienter at the pleading stage, the certification requirement alone is not quite the veil piercing tool it was originally thought to be.

Conclusion

While post-Enron legislation and regulation, namely Sarbanes-Oxley, was intended to change the complexion of professional liability for corporate officers and accountants alike, in reality, there hasn't been a major shift in claim activity.[19] In fact, rather than rectify the crisis of confidence in the U.S. corporate governance and financial reporting, Sarbanes-Oxley has actually created a number of gray areas in its interpretation which has led to legal experts, auditors, and corporate executives searching for some solid guidance and alternative solutions.[20] Moreover, it appears that even when a corporate officer submits a fraudulent certification to the SEC, that the false certification alone will not be sufficient to establish scienter and to pierce the corporate veil. Despite the codification of personal liability provisions such as §§ 302 and 906 of SOX, it appears that litigants must still leap the high hurdle of demonstrating a strong inference of scienter that was established long before SOX in order to pierce the corporate veil and hold those officers who committed fraud responsible for the devastation their actions have caused.

REFERENCES

- [1] Sarbanes Oxley Act of 2002, Pub. L. No. 107-204, 116 Stat. 745 (2002) (codified in various sections of the U.S.C.).
- [2] Bill Carlino, *After Major Repairs, Profession Awaits Detail Work*, Accounting Today, Sept. 20, 2004, at 42.
- [3] Jonathan D. Glater, *New Rules Make it Easier to Charge Executives, But Not to Send Them to Prison*, New York Times, July 5, 2005, at 5.
- [4] Franklin A. Gevurtz, *Piercing Piercing: An Attempt to Lift the Veil of Confusion Surrounding the Doctrine of Piercing the Corporate Veil*, 76 Or. L. Rev. 853 (Winter 1997).
- [5] *See* Schattner v. Girard, Inc., 668 F.2d 1366, 1370 (D.C. Cir. 1981).
- [6] 15 U.S.C. § 78u-4(b)(2) (emphasis added).
- [7] *In re* Comshare, Inc. Sec. Litigation, 183 F.3d 542, 548 (6th Cir. 1999).
- [8] *See* Helwig v. Vencor, Inc., 251 F.3d 540, 551 (6th Cir. 2001).
- [9] *See* H.R. Conf. Rep. No. 104-369 (1995).
- [10] *See* Acito v. IMCERA Group, Inc., 47 F.3d 47, 52 (2d Cir. 1995); *See* Novak v. Kasaks, 216 F.3d 300, 307-08 (2d Cir. 2000), *cert. denied*, 531 U.S. 1012 (2000).
- [11] *See* *In re* Silicon Graphics Litigation, 183 F.3d 970, 974 (9th Cir. 1999).
- [12] *See* Bryant v. Avado Brands, Inc., 187 F.3d 1271, 1287 (11th Cir. 1999); *see also* Garfield v. NDC Health Corp., 466 F.3d 1255, 1265-66 (11th Cir. 2006)
- [13] *See* Tellabs, Inc. v. Makor Issues & Rights, Ltd., 127 S. Ct. 2499, 168 L. Ed. 2d 179 (2007).
- [14] *See* *In re* Lattice Semiconductor Corp, Securities Litigation, 2006 U.S. Dist. LEXIS 262 (D. Or. 2006).
- [15] *See* Frank v. Dana Corp., 2007 U.S. Dist. LEXIS 61307 (N.D. Ohio 2007).
- [16] *See* *In re* Silicon Storage Technology, Inc. Securities Litigation, 2007 U.S. Dist. LEXIS 21953 (N.D. Calif. 2007).
- [17] *See* Srebnik v. Dean, 2006 U.S. Dist. LEXIS 73836 (Colo. 2006).
- [18] *See* *In re* OCA, Inc. Securities and Derivative Litigation, 2006 U.S. Dist. LEXIS 90854, at *72 (D. Louisiana 2006) (citing to *In re* Lattice Semiconductor Corp. Sec. Litig., 2006 U.S. Dist. LEXIS 262 (D. Or. 2006)).
- [19] Roger Russell, *Facing the Storm: Planning and SOX Worrying Insurers*, Accounting Today, July 26, 2004 – August 8, 2004 (quoting Ken Mackunis, president of Aon Insurance Services, the administrator for the American Institute of CPAs Professional Liability Insurance Program).
- [20] *See* Dr. Sabyasachi Ghoshray, *Impact of Sarbanes-Oxley on Multiple Listed Corporations: Conflicts in Comparative Corporate Laws and Possible Remedies*, 10 ILSA J. Int'l & Comp. L. 447 (Spring 2004).

COMMUNITY EMPOWERMENT AND SERVICE LEARNING PRACTICES THROUGH INFORMATION SYSTEMS AND TECHNOLOGIES IN NEW YORK CITY PANEL

James Lawler, Pace University, 163 William Street, New York, New York 10038, 212-346-1013, lawlerj@aol.com

PARTICIPANTS

Jean Coppola, Pace University
Jonathan Hill, Pace University,
James Lawler, Pace University
Pauline Mosley, Pace University
Andrea Taylor, Pace University

SESSION FORMAT

Service learning continues to be an evolving but exciting field in academia. The learning engages students in civic practices that empower community non-profit organizations in real solutions. More than half of higher level institutions in the United States have service learning in their curricula. From this panel, instructors of a major institution in the northeast demonstrate practices of service learning in computer science and information systems curricula that are helping community organizations on solutions which are taking immediate advantage of 21st century technologies, including those of the World Wide Web. This panel is an introduction to *best of class* pedagogy in service learning at one of the pioneers in service learning in northeast institutions.

The focus of the panel is on projects that seamlessly include technology and service learning in the computer science and information systems curricula of Pace University in downtown New York City and in Brooklyn, New York. The projects demonstrated at the panel are of agile methodologies of community focused social and technical processes that integrate disadvantaged citizens of communities and undergraduate students of an institution in solutions of technology. This panel is focused on the benefits for students and for local non-profit organizations of integrating technology and the Web in a service learning strategy.

Service learning is based on the partnership of a Center for Community Outreach, which is presented at the panel. Such centers are a critical success factor in identification of non-profit organizations desiring help and in liaison of the organizations with the computer science and information systems schools and other schools of an institution. Funding of innovation in curricula design and in empowering practices in service learning is facilitated by partnerships with a center of outreach resident in an academic institution. The panel furnishes a practical pedagogy and partnership strategy that may be applied by instructors considering formal inclusion of service learning in the curricula of their institutions. The panel is apt for a conference in metropolitan New York City.

Something Else They Didn't Teach Us: Management Education for Non-Traditional Managers. The Case of Library Directors

PANEL THEME – ABSTRACT

This panel theme has evolved from a broader research agenda on defining and understanding the process of educating the individuals who hold management positions in non-traditional businesses. Examples of such non-traditional business managers, and the graduate degrees they hold, include: Library Directors (MLS), Pastors (M.Div), Primary and Secondary School Principals (Ed.D), Museum Directors (MFA), College Provosts and Presidents (Ph.D.) and Physicians (M.D.).

The academic preparation for an individual earning a discipline-specific degree, as described above, is fundamentally different from that taken by the individual who has sought an advanced degree in management with the intention of assuming a leadership role within *any* industry. The Master of Business Administration (MBA) degree is the terminal degree that prepares an individual to be a manager regardless of discipline.

The Library Director position is the prevalent managerial position within the library profession. Librarians work in a broad range of institutions and perform a broad range of functions; in the performance of these multitudinous functions there exists the role of management. Library directors must deal with all the personnel decisions related to full-time, hourly and student employees. Depending on the organizational structure of their institutions, they must work with union employees and/or non-union employees. They must be aware of the range of human resource related topics, such as reasonable accommodation, sexual harassment, fair labor standards, equal employment opportunity provisions, and more. (Mackenzie and Smith 2007).

Where do library directors, and the librarians who perform various management functions as part of their work, receive their management training? A review of the curricula of 48 MLIS programs accredited by the American Library Association (ALA) revealed that, for the most part, library managers receive their management training while on the job. Of the programs reviewed in this study, 43.8% did not require a management-related course within its degree requirements. And 81.3% of the programs did not require an internship. (Mackenzie and Smith 2007)

An examination of graduate level course syllabi within the ALA accredited library programs revealed that only 58.3% of the management courses included human resource management and only 54.2% included strategy, planning and process. This suggests that the library profession has not yet agreed upon the requirements for preparing future librarians for managerial positions. (Mackenzie and Smith 2007). The results of a recent empirical study revealed that 55.1% of library directors surveyed stated that graduate library school did *not* properly prepare them to be library directors.

A review of the recently published professional literature suggests that the library profession is not alone in its concern regarding a lack of management training in its master-level programs. Problems with Master of Fine Arts (MFA) programs, Master of Social Work (MSW) programs, and wildlife student programs, have all been topics of articles. All the authors share a common concern; that newly educated professionals lack adequate training in subject areas involving crucial management skills. (Nesoff 2007, Kroll 2007, Rhine 2007).

There has been little specifically written on the professional development of library directors. Journal articles and monographs found on this topic primarily originate with the ALA. Browsing the ALA Editions publications collection reveals a subject classification of "Administration and Management." However, most of the titles in this class refer to the management of technology,

collections, customers, funding agencies, trustees, assessment demands, and intellectual property. The focus on the management of people, as a traditional manager, is a sub-focus even within this venue.

There are, encouragingly, relevant titles that support the view that library directors are traditional managers and require skill and knowledge development in that arena (Giesecke 2001, Montgomery & Cook 2005, Weingand 2001.) Similar is the viewpoint that the information field has been dissatisfied with the education emerging from library and information science programs (Koenig 2007, Nichols and Koenig 2005).

The results of recent research will introduce the panel topic to the audience. The focus of the panelists will be on the following topics: Management education, Library School education, the role of traditional managers, and the role of library directors.

The binding theme is the academic preparation of individuals who choose to enter the field of librarianship and in turn assume *management positions*.

The panel session will begin with a brief description of the panel's overall theme and its relevance and importance to the broader NEDSI community. This will set the stage and provide context to the individual panelist backgrounds. The panel session will combine a presentation of the cited studies, but the focus will be the discussion among the panelists and the audience. The expertise of this panel draws together the needs of the employer, the needs of the non-traditional manager, and the providers of academic preparation. The panelists offer varying viewpoints and perspectives into this topic.

The panel includes the following experts:

Dr. Maureen Mackenzie

Areas of expertise: MBA Education

Affiliation: Dowling College, Management education

Dr. James P. Smith

Areas of expertise: Library Management

Affiliation: St. Francis College. Director of the St. Francis College Library

Dr. Michael Koenig

Area of expertise: Library School Education

Affiliation: Palmer School of Library Science, Long Island University

Ms. Lauren Spatig, MBA

Area of expertise: Corporate Management, Employer needs

Affiliation: Allstate Insurance Company and Fielding Graduate Institute

WORKS CITED

Giesecke, Joan. (2001). *Practical Strategies for Library Managers*. ALA Editions.

Koenig, Michael E. D. (2007, March). Looking back: A review of four key articles from Volume One. *Education for Information – 25th Anniversary Issue*.

Kroll, A. J. (2007). Integrating professional skills in wildlife student education. *Journal of Wildlife Management*, 71(1), 226-230, February 1. URL: <http://www.proquest.com>. Last viewed December 21, 2007).

- Mackenzie, Maureen L., and Smith, Jim P., (2007). An exploratory study of libraries and their managers: Management education for Leaders of non-traditional businesses. *Proceedings of the 70th Annual Meeting of the American Society for Information Science and Technology*, October 2007, volume 44. Published on CDROM. ISBN: 0-87715-539-9. Pages 960-975.
- Montgomery, Jack G., and Cook, Eleanor I. (2005). *Conflict Management for Libraries*. ALA Editions.
- Nesoff, Irwin. (2007). The importance of revitalizing management education for social workers. *Social Work*, 52(3), 283-5, July 1. URL: <http://www.proquest.com>. (Last accessed December 21, 2007).
- Nichols, Jerry, and Koenig, Michael. (2005.). Developments in post-masters education for public library administration in North America: With particular reference to the Public Library Administrator's Certificate Program at The Palmer School of Library and Information Science. IN Genoni, Paul and Graham Walton (eds). *Continuing Professional Development- Preparing for New Roles in Libraries: A Voyage of Discovery*. IFLA Publication 116. ISBN 10: 3-598-21844-3192-197. Abstract published online at <http://www.ifla.org/IV/ifla71/programme2005-cpdwl-e.pdf>, p.16 (2005)
- Rhine, Anthony S. (2007). The MFA in theater management and the MBA: An examination of perspectives of decision makers at theaters in the United States. *Journal of Arts Management, Law, and Society*, 37(2), July 1: 113-121,123,125-126. URL: <http://www.proquest.com> (Last accessed December 21, 2007).
- Weingand, Darlene E. (2001). *Administration of the Small Public Library, 4th Edition*. ALA Editions.

A STUDY ON IMPROVEMENTS OF COLLEGE STUDENTS IN QUANTITATIVE APTITUDE

Jinchang Wang, jinchang.wang@stockton.edu
Gurprit Chhatwal, gurprit.chhatwal@stockton.edu
Shaoping Zhao, shaoping.zhao@stockton.edu
School of Business
Richard Stockton College of New Jersey
Jimmie Leeds Rd., Pomona NJ 08240, USA

ABSTRACT

Quantitative reasoning is a capstone capability that college students need to acquire for dealing with the challenges they will face in the real world. This paper presents a study that assesses students' improvement in quantitative aptitude during their college years. We used the convenient "SAT sample assessment approach", on a sample of 120 business students majoring business, to compare their current status with that when they entered college. The results showed students' improvement during their college years. We further studied the correlations between students' improvement and variables such as GPA, credit hours earned, and "math anxiety". The results posed interesting questions on patterns and causes of student's improvement. We present the implications of our results and propose further studies.

Key Words: outcomes assessment, quantitative aptitude, college education, statistical analysis

I. INTRODUCTION

Quantitative reasoning is a capstone capability that higher institutions expect their students to possess to deal with the challenges of this dynamic electronic age. Although calculators relieved us from irritations of numerical computing, ubiquitous computers have brought in a host of requirements on logical and quantitative reasoning to effectively use it and to develop more applications. Almost all colleges put quantitative reasoning aptitude as one of the fundamental capabilities students need to acquire during their college years. To the chagrin of educators in higher institutions, that capability is lacking in many of them at graduation [7]. Data from the Department of Education shows that American literacy progress has stagnated during the past several years [5]. Some recent studies attempted to explore the underlying reasons for this state of affairs. Dar-Nimrod presented the controversy on whether math ability was nature or nurture [2]. Ruffins addressed "math anxiety" as the worst enemy of studying math and the ways to overcome it [4]. Jones and Byrnes looked at influence of students' characteristics on studying math [3].

Many institutions have moved, or are moving, their missions in the direction of improving and enhancing the student's quantitative reasoning capability for the challenges they will face in their careers. Assessing students' existing capabilities and examining the efficacy of any program of study are the first steps in initiating a curriculum improvement project. Good outcomes assessment activities meaningfully inform the administration, instructors and students about the strengths, weaknesses, and effectiveness of the learning process. This information is useful for curriculum setting and pedagogical improvements to enhance students' learning.

In this paper, we present a study of quantitative aptitude assessment for business students at Richard Stockton College of New Jersey. The purpose of the study is to investigate whether students' quantitative

reasoning capability is improved during their study in college, and the factors that cause the improvement, if any. We used a shortened SAT test as the assessment instrument, and compared a student's assessment test score with his/her SAT score. We studied the correlations between improvements in students' performance and variables such as GPA and credit hours earned. We also studied whether "math anxiety" existed with our students and how it affected students' progress in mathematics learning.

The study showed that in general students exhibited an improvement in their quantitative reasoning aptitude during their college years, and the students whose SAT scores were lower tended to have greater improvement. They showed that students' improvement was not correlated significantly to the expected causal factors such as courses taken, GPA, and transfer hours. We also found that the improvement is not significantly correlated to student's attitude towards math. Although the results are kind of unexpected and seem confusing, they are inferred facts from the data. Actually, they pose a series of interesting questions that lead us to do further research.

II. THE SAT SAMPLING APPROACH

The SAT sampling approach was developed by Wang *et al* for efficiently assessing quantitative aptitude of college students [6]. It is a special paired-test assessment approach in which the "pre-test" is SAT that a majority of college students take before entering college. Only one assessment test is needed as the "post-test".

The assessment test in the SAT sampling approach is actually a shortened SAT test, which is composed of questions randomly selected from the quantitative part of old SAT exams, but much shorter than the SAT test. There are usually 60 quantitative questions in a SAT test, which takes 75 minutes (1.25 minutes per question) for students to do them. Wang *et al* argued that a test of as few as 20 questions is statistically comparable to the SAT quantitative test of 60 questions [6]. A 25-minute test of 20 questions can be conveniently and flexibly administered at any time of a semester in most classes. Four key steps in the SAT sampling approach are:

1. Construct the assessment test that contains 20 questions selected from the quantitative part of past SAT exams.
2. Select students of some classes to take the assessment test.
3. Collect SAT data for those students.
4. Statistically compare the assessment test results with the participating students' SAT scores to evaluate improvement in their quantitative capability.

III. ADMINISTERING THE ASSESSMENT TEST

In our study for students' improvements in quantitative aptitude, we used the SAT sampling approach. We needed to administer only one assessment test to students, since their scores of "pre-test", SAT, were available from school's database.

We selected 20 questions for our assessment test from past ten SAT tests using 'random-within-category' method. We reviewed 600 quantitative questions in ten past SAT exams [1], and marked each question with its subject area such as algebra or geometry. We found that in a SAT test there were approximately 30% of arithmetic questions, 30% of algebra, 30% of geometry, and 10% from other miscellaneous subjects such as probability, data interpretation, and logic. We maintained the same proportion among the four subject areas in our 20-question assessment test. For the subject of algebra, for example, we selected six questions at random from the SAT questions classified as "algebra". Our assessment test eventually

consisted of six questions in algebra, six in geometry, six in arithmetic, and 2 miscellaneous. The assessment test is attached as an appendix.

To be comparable to SAT score, our assessment test score must be converted to an 800-point-base. We did the conversion as shown below.

Step 1: Calculating *Points-Earned* for a student

$$\text{Points-Earned} = R \cdot (C - 0.25I)$$

where C=number of correct responses,
 I=number of incorrect responses,
 $R = (\text{number of questions in SAT}) / (\text{number of questions in our test})$
 $= 60/20 = 3$

Step 2: Converting *Points-Earned* into a score between 200 and 800 by using the SAT Conversion Table available in SAT exam books (The College Board, 2000).

We included three survey questions in the assessment test to survey students' mathematical dispositions. Students chose from 'strongly agree' to 'strongly disagree' for the following three statements:

Q1: Mathematics is a challenging subject.

Q2: Mathematics is a useful tool in one's career and daily life.

Q3: I feel more comfortable on mathematics now than before.

The assessment test was administered to business students enrolled in a required course at the end of spring and fall semesters in 2006. The course was Quantitative Business Methods. Extra credit was given to stimulate students' interest and to ensure their seriousness in taking the test. Two hundred and four (204) students took the test. Valid SAT/ACT scores were available for 112 students (53 in spring and 59 in fall). Our statistical analyses are for the 112 students who had both the assessment scores and either the SAT data or ACT data available. Of the 112 students, 60 students did not put their names on the survey sheet. Therefore, we only have 52 students' data for the analyses regarding the three survey questions.

IV. TEST RESULTS AND STATISTICAL ANALYSIS

In this section, we present the summarized result of the assessment test, and its comparison with the students' previous SAT scores. We also show the outcomes of the regression analysis among various factors.

Table 1 shows the descriptive statistics of the assessment test and the SAT quantitative test for the 112 students in our study.

Table 5.1 Descriptive Statistics Comparison (112 students)

Group	No. of students	Assessment Test (800 base)				SAT Quant. part (800 base)		
		Test time	Median	Avg.	Standard deviation	Median	Avg.	Standard deviation
A (Spring)	53	Spring 2006	640	633.21	72.95	570	552.5	72.29
B (Fall)	59	Fall 2006	640	632.54	90.89	520	518.4	78.64

We further studied the correlations between score-improvement and the respective SAT score / assessment score. Furthermore, to identify the factors that may have contributed to an improvement in

scores, we looked at the relationship between score-improvement and GPA, total credit hours taken, total quantitative courses taken, and total transfer hours. We used the regression analysis in Excel to explore the correlations. Table 2 provides the results.

Table 2. Correlations between Improvement and Other Factors (Sample size: 112)

Variable Y	Variable X	p-value	Positive or Negative relation	Significant or not
Assessment Test	SAT	1.14 E-06	+	Sig.
Assessment Test	GPA	0.007	+	Sig.
Score-improvement	Assessment test	6.4294 E-10	+	Sig.
Score-improvement	SAT	2.388 E-08	-	Sig.
Score-improvement	Quantitative hours taken	0.70152	+	Not
Score-improvement	GPA	0.816	-	Not
Score-improvement	Total credit hours taken	0.2315	-	Not
Score-improvement	Total transfer hours	0.559	-	Not

Table 3 shows correlations that involves the three survey questions

Q1: Mathematics is a challenging subject.

Q2: Mathematics is a useful tool in one's career and daily life.

Q3: I feel more comfortable on mathematics now than before.

Student selected one of the five choices for each questions: 1 for strongly agree, 2 for agree, 3 for neutral, 4 for disagree, 5 for strongly disagree.

Table 3. Correlations with respect to three survey questions (Sample size: 52)

Variable Y	Variable X	p-value	Positive or Negative relation	Significant or not
Assessment Test	Q1	0.009	+	Sig.
Assessment Test	Q2	0.3905	-	Not
Assessment Test	Q3	0.5573	-	Not
Score-improvement	Q1	0.6398	+	Not
Score-improvement	Q2	0.2665	-	Not
Score-improvement	Q3	0.7044	+	Not

V. DISCUSSIONS

It would be logical and reasonable to assume that "SAT score" reflects students' quantitative aptitude upon entering college, "assessment test score" represents their current aptitude level, and "score-improvement" indicates the extent of improvement during their college years.

Results of our study show that students' did improve their quantitative reasoning aptitude during college years, and that improvement was significant. Students whose SAT scores were lower tended to demonstrate greater improvement. The results of statistical regressions were interesting. They show that score-improvement is not significantly correlated to the expected causal factors such as courses taken, GPA, and transfer hours. And score-improvement is not significantly correlated to student's attitude

towards math. These unexpected weird-looking results pose interesting questions for us to discuss here and to investigate further research.

(1) The study shows that students' quantitative aptitude did improve by an average of almost 10% after joining the college (see Table 1).

Note that this result on student's significant improvement in math is not inconsistent with the 2003 US Department of Education study which showed that mathematic literacy progress had stagnated during the past several years (US Department of Education, 2003). DOE was comparing student groups year by year. Student group in 2003, for example, was different from that in 1999. But our study was comparing performances of same students in different years.

(2) Student's score-improvement is significantly correlated to both the assessment test score and the SAT score (see Table 2). It poses a positive correlation to the assessment score, while a negative correlation to the SAT score. The implication is that the students who did poorly in math before entering college or who are doing well in math currently tend to demonstrate more improvement.

(3) Student's improvement is not significantly correlated with GPA, total credit hours taken, total quantitative hours taken, or total transfer hours (see Table 2). Also, score-improvement is not significantly correlated with student's attitude towards math which is represented in the three survey questions (see Table 3).

These "non-correlations" look unexpected. They seem to suggest that mathematical courses in college have no effect on student's improvement in quantitative aptitude. However, the results in our study are not inconsistent with those in US Department of Education's 2003 study which indicated that American students' mathematic literacy progress had stagnated during the past several years (US Department of Education, 2003-a). How could students' math aptitude remain "stagnated" in the face of billions of dollars the government spent on higher education, especially the monies spent on projects to specially improve students' quantitative reasoning ability? A first reaction may conclude that improvement in quantitative aptitude is not related to the federal money spent for improving such aptitude. In other words, the money that the government spent for improving American students' quantitative aptitude was ineffective. But it is too soon for us to theorize that the efforts and money US DOS invested was totally in vain based on the study results only. That is because there are many other factors affecting student's math literacy. There has been no evidence to show that in absence of the government funding, quantitative aptitude could have been able to maintain "stagnated" without going down during the time period in question. Similar logic applies in our case. Those statistical "non-correlations" in the results of our study are not sufficient evidences to automatically theorize that our curriculums were ineffective. There are a myriad of factors that impact the progress of a student's quantitative aptitude, many of which may have been missed in our study. We need to further explore those causal factors that affect student's improvement in their quantitative aptitude.

The "non-correlations" observed in our study indicate complexity of quantitative aptitude. Quantitative aptitude is not just quantitative knowledge. It is the sense of math. It is the way of thinking logically and mathematically. Quantitative aptitude is not as simple as, so to speak, "it grows so long as we water it". Those "non-correlation" results take us to the unsolved issue of characteristics of quantitative aptitude: - Is quantitative aptitude nature or nurture, or somewhere in between?

It looks likely that math aptitude is similar to music aptitude. The sense of music could be improved by education, but only in a limited way. There are many unknowns about aptitudes of music and math. For example, how much of such aptitude, and in what aspects, can be acquired through education; whether the

aptitude progresses continuously or intermittently; what is the most effective pedagogical way to enhance student's sense of quantitative and logic reasoning.

(4) The “non-correlation” between student's improvement and total number of credit hours transferred in from other colleges (see Table 2) implies that other schools are as good, or as bad, as we are in improving student's quantitative aptitude.

(5) Regression analysis indicates that student's assessment score is positively and significantly correlated with the response to Q1 - “Mathematics is a challenging subject”, where “strongly agree” is represented by 1, and “strongly disagree” by 5 (see Table 3). This is not very surprising as it indicates more comfortable a student is with math, better quantitative aptitude he/she possesses currently.

(6) Students' assessment test scores are strongly and positively correlated with their SAT scores (see Table 2), which indicates that our assessment test is a valid one, or as good as SAT, in testing students' quantitative aptitude.

REFERENCES

- [1] The College Board, 10 Real SATs, 2nd Edition, Henry Holt and Company, LLC. Editor Cathy Claman
- [2] Dar-Nimrod, I. Math ability in women - nature versus nurture. *Nanotoday*, Vol.2 No.3, June 2007, p.56-57
- [3] Jones, K. & Byrnes, J. Characteristics of students who benefit from high-quality mathematics instruction. *Contemporary Educational Psychology*, Vol.31 No. 3, July 2006, p.328-343
- [4] Ruffins, P. A real fear. *Diverse Issues in Higher Education*, Vol. 24 No.2, 17-19, Mr 8 2007
- [5] US Department of Education. National Assessment of Adult Literacy, 2003, <http://www.ed.gov>.
- [6] Wang, J, Chhatwal, G. & Zhao, S. The SAT sampling approach for assessment of quantitative aptitude of college students, *Proceedings of DSI Annual Meeting, San Antonio, TX, Nov. 2006*, p.26871-26875
- [7] Wilkins, J. L. Preparing for the 21st Century: The Status of quantitative Literacy in the United States. *School Science and Mathematics*, Vol. 100(8).2000, p.405-418

PAY TO PLAY: A REPORT ON THE USE OF A WEB-BASED, PAY-FOR-USE OPERATIONS MANAGEMENT SIMULATION

Thomas W. Sloan, University of Massachusetts Lowell, (978) 934-2857, Thomas_Sloan@uml.edu
David A. Lewis, University of Massachusetts Lowell, (978) 934-2758, David_Lewis@uml.edu

ABSTRACT

The value of experiential learning is well documented, and online games and simulations are an efficient and effective way to convey essential concepts in decision sciences and operations management. This paper reports on the use of a web-based, pay-for-use simulation called Littlefield Technologies, where student teams compete by managing factories that assemble digital satellite receivers. We focus on two questions. First, does the simulation improve learning outcomes? And second, do the students find the simulation valuable? Survey results indicate that while students feel that they learn a great deal from the simulation, they are willing to pay little, if anything, for it. We discuss how these insights can be used to improve decision sciences education.

Key words: Experiential learning, games and simulations.

INTRODUCTION

All students enrolled in the College of Management at the University of Massachusetts Lowell (UMass Lowell) are required to take an introductory course in Operations Management (OM). A key challenge in management education in general — and OM education in particular — is connecting theory to practice. Experiential learning has proven to be a valuable tool in addressing this challenge, and simulations and games have become a popular way to improve management education by engaging students in ways that textbooks, homework problems, and case studies can not [1] [2].

How are instructors to create effective, real-world experiences for students in their operations management and decision sciences courses? Advances in technology have created unprecedented opportunities for teachers to move beyond simple games based on dice, tinker toys, and construction paper. In particular, many computer-based games and simulations have emerged in recent years to help teach students basic concepts related to decision sciences, operations, and supply chain management. The simulations vary greatly in terms of complexity, ability to capture system dynamics, and cost. In this paper, we report our experience using Littlefield Technologies, a web-based factory simulation, in an undergraduate operations management course.

BACKGROUND

Experiential Learning

A major challenge in conveying the principles of operations management is the students' lack of experience. For example, while all students have seen and operated DVD players, few have seen

one being assembled, and therefore they have a difficult time imagining the systems and processes needed to manufacture a DVD player successfully. Experiential learning — involving role play, games, and simulations — has long been recognized as an effective tool in enhancing the educational process (see [4] or [5], for example). This is achieved by actively engaging students, enabling them to make their own discoveries in a “safe” environment, and providing rapid feedback on decisions and performance. Beyond teaching classroom concepts, experiential learning can also develop students’ ability to work in teams, an essential skill in today’s business environment [3].

Online Games and Simulations

Simulations enable students to experience most dimensions of a business operation on a small scale and in accelerated time. Not surprisingly, simulations have become a mainstay of many management courses [2] and are particularly well-suited to OM courses. Perhaps the best-known OM-oriented simulation is the Beer Game, developed at MIT [7]. For a discussion of different types of OM games and simulations, refer to [9].

Advances in computer technology have made simulations easier to administer by automating some of the “bookkeeping” tasks and facilitating synchronization of various aspects of games. Evolution of the world wide web has created another avenue to expand and refine simulations. Online games enable student access from virtually any computer and further reduce the administrative burden for instructors. Wood [8] reviews many of the simulations available for OM and related topics, most of which are computer-based. Some are played on stand-alone computers, while others are played on-line. Some games last only 15 minutes, while others last for days or weeks. A few simulations are free; many cost \$5 to \$15 per student, and others cost \$45 or more per student. The scope of decision making ranges from narrow (e.g., determining optimal order quantity for a single factory) to broad (e.g., managing an entire supply chain from product development through distribution).

Student Background

The University of Massachusetts Lowell is located about 25 miles northwest of Boston along the Merrimack River. The university has about 11,000 undergraduate students, most of whom are Massachusetts residents. Like many students attending urban, public schools, our students struggle to make ends meet. Many students work over 20 hours per week to pay for tuition and other expenses, and approximately 73 percent of undergraduate students receive some financial assistance. A key motivation for this study was the attempt to maximize educational value: How can we provide the most effective education possible without placing an undue financial burden on the students?

Littlefield Technologies

Littlefield Technologies is a web-based factory simulation developed and administered by Responsive Learning Technologies (Responsive). The factory assembles digital satellite system receivers using a three-station, four-step process. Students are divided into groups of four and make decisions about how to operate the factory with the goal of maximizing profit. Teams can monitor inventory levels, equipment utilization, order arrivals, etc. to help guide their decision making. In addition, each team can view its performance relative to the other teams as well as a “do-nothing” team, which makes no changes to the initial settings. The demand varies over time; however, the mean demand increases steadily at the beginning of the simulation and then levels

off. (The students are not informed of this pattern, but may observe it by studying the order arrivals.) We used a two-phase assignment, which is typical. The first phase, which we ran in week six of the course, focuses on capacity, and the only actions available to the students are addition or subtraction of machines at each station and the job sequencing priority at one station. The first phase lasted for one week, which is equivalent to 268 simulated days.

The second phase, which we ran in week 12 of the course, and also lasted one week, focuses on responsiveness. In addition to adding machines and changing the queue discipline, teams can make changes to the raw material order quantity and reorder point in this phase. Teams can also choose among three different sales contracts, earning more revenue for quicker turnaround times.

The instructor can change virtually any of the parameters: demand distribution, length of simulation, speed of simulation, etc. See [6] or [9] for more details about the default scenario as well as a discussion of other typical scenarios.

The two-part Littlefield Technologies simulation normally costs \$15 per student. Registration codes for the simulation can be purchased directly from Responsive or can be offered through the campus bookstore (usually with some mark-up). Other payment options are available. Given our lack of experience with the game, we were reluctant to have students pay this fee on top of the textbook cost (which retails for approximately \$150). Fortunately, we were able to get funding for the simulation by receiving a university grant for innovative teaching. This enabled us to learn about and experiment with the simulation without creating a financial burden for the students. Unfortunately, we did not receive notice of the availability of funds until the fifth week of class, which did not allow for the coordination of lecture material with the first phase of the simulation.

DATA AND ANALYSIS

Methodology

The simulation was used in two sections of an introductory operations management course, with a total of 68 students participating. After the first phase of the simulation, students were given a nine-item, online survey regarding their experiences. Fifty seven students responded to the survey. After the second phase of the simulation, students were surveyed again, with two additional questions added to the survey, and 52 students responded. The complete list of survey questions is available upon request from the authors.

Results

We feel that the two most relevant items in the survey are:

- This game contributed to my understanding of capacity management and inventory management.
- How much would you be willing to pay [as an additional fee] to have access to this simulation in class?

A five-point Likert scale was used for the first item, with possible responses of “strongly disagree” (SD), “disagree” (D), “neither agree nor disagree” (N), “agree” (A), and “strongly agree” (SA). The results for this item are displayed in Figure 1. After the first round, approximately 80 percent of the students responded positively to the first item (either “agree” or “strongly agree”). That number rose to approximately 92 percent after the second round.

Furthermore, the number responding with “strongly agree” rose by more than 75 percent from the first round to the second.

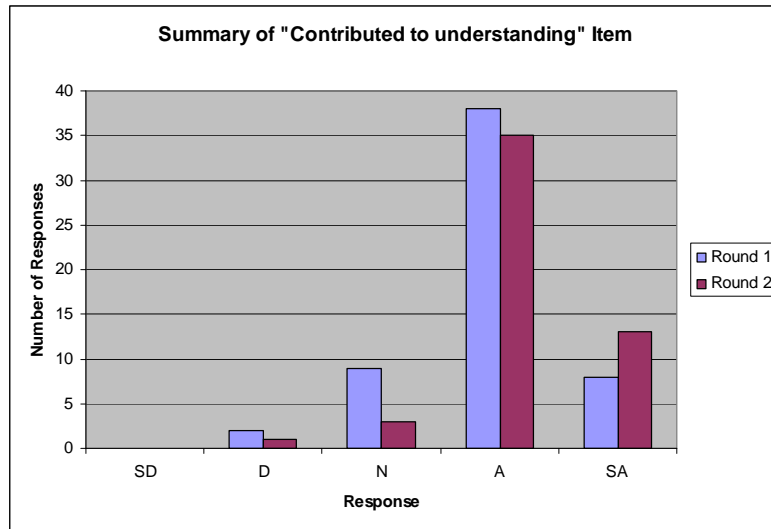


FIGURE 1

Possible responses for the second item changed from the first round to the second. In the first survey, the possible responses were “\$0, no value,” “\$0, has value, but not at extra cost,” “up to \$10,” “up to \$15,” and “up to \$20.” Based on the observed results, we modified the second survey to allow more choices; possible responses were “\$0, no value,” “\$0, has value, but not at extra cost,” “up to \$5,” “up to \$10,” “up to \$15,” and “more than \$15.” The results for the “Willing to pay” item are displayed in Figure 2. After the first round, approximately 70 percent of the students were not willing to pay for the simulation, although they found it valuable. Nearly 25 percent were willing to pay something extra, while about 5 percent saw no value in the exercise. After the second round, however, roughly 36 percent of respondents were willing to pay something extra for the simulation, and no one felt that it had no value.

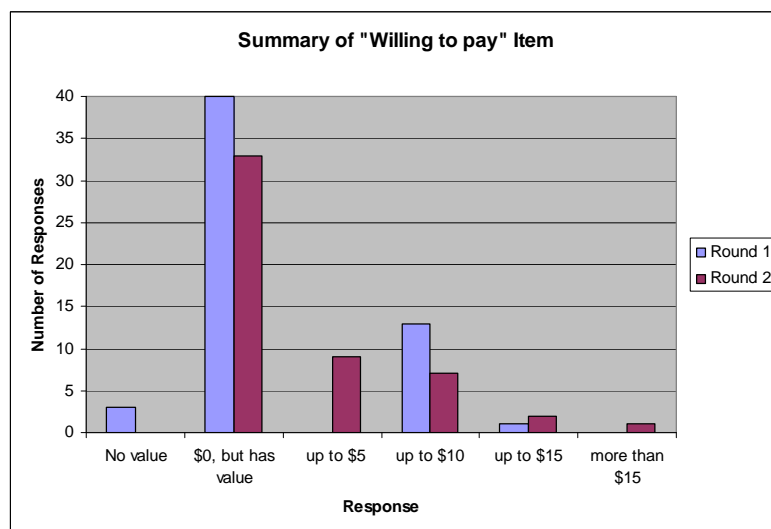


FIGURE 2

Interestingly, about 56 percent of respondents who selected “agree” or “strongly agree” for the “Contributed to understanding” item also selected “\$0, has value, but not at extra cost” for the “Willing to pay item.” Put differently, the students find the simulation valuable but would rather not pay for it. Although we have no concrete data to support it, we conjecture that they might have similar feelings regarding textbooks, tuition, and professors’ salaries.

SUMMARY AND CONCLUSIONS

We found the simulation to be useful and learned some important lessons along the way, some of which include:

- It is essential that some part of the assignment be for a grade. As recommended by Responsive, we placed 20 percent of the weight on the performance of the teams; this encourages the students to take the assignment seriously. However, most of the grade weight was on the post-game write-ups.
- There was some confusion and anxiety during the first round. Students were not sure how the simulation worked, and their post-simulation comments indicated that they wanted more instruction about how often to check results, when to update decisions, what things they should be looking for, etc.
- Much of the anxiety mentioned above was eliminated in the second round, as they had a much better idea about what to expect. Thus, we feel that a two-round simulation is of great benefit. In addition, we would consider placing more grade weight on the second round than on the first.
- Since the simulation is designed to be run entirely out of class, there was some concern about communicating with group members and coordinating decisions. Some teams clearly did not have a coherent plan for how and when decisions would be made and suffered from a “too many cooks spoil the broth” effect.

We draw two main conclusions from our experience with Littlefield Technologies. First, the simulation was very positive, and Responsive was very helpful in setting up and running the simulation. As we expected, the students really liked the idea of competition. It helped get them engaged and forced them to think about the material more actively. Ideally, we would be able to directly test whether the game improved learning, but there are several obstacles to doing so, such as the difficulty of assessing knowledge prior to the simulation and the difficulty of having a control group. Although it is clearly not the same direct assessment, students reported that the simulation did increase their understanding of key operations management concepts.

Our second conclusion is that while students find the simulation valuable, they do not want to pay for it. This result is not surprising given the increasing financial pressures that students face and the increasing reliance on course supplements. The real issue is: What, if anything, should instructors do about it? Do we have an obligation to seek lower-cost alternatives? Is there a way to reduce other costs, for example by choosing a lower-cost textbook? Or is it simply a matter of “marketing” — i.e., informing students of the benefits and trying to persuade them that it is worth the cost?

Fortunately, we have received enough financial support to fund another semester of the simulation (in spring 2008). We hope to make good use of the lessons learned and expect to collect more data which will allow us to more objectively and rigorously examine learning outcomes.

ACKNOWLEDGEMENT

The authors gratefully acknowledge funding from the UMass Lowell Council on Teaching, Learning, and Research in the form of an “Explorations in Teaching and Learning” grant.

REFERENCES

- [1] Ben-Zvi, T., Carton, T. C. “From rhetoric to reality: Business games as educational tools,” *INFORMS Transactions on Education*, 2007, 8(1), <http://ite.pubs.informs.org/Vol8No1/Ben-ZviCarton/>.
- [2] Faria, A. J. “Business simulation games: Current usage levels — an update,” *Simulation and Gaming*, 1998, 29(3), 295–308.
- [3] Kayes, A. B., Kayes, D. C., Kolb, D. A. “Experiential learning in teams,” *Simulation and Gaming*, 2005, 36(3), 330–354.
- [4] Kolb, D. A. *Experiential Learning: Experience as the Source of Learning and Development*. Englewood Cliffs, NJ, Prentice Hall, 1984.
- [5] Kraft, D., Sakofs, M. (Eds.) *The Theory of Experiential Education*. Boulder, CO, Association for Experiential Education, 1984.
- [6] Miyaoka, J. “Making operations management fun: Littlefield Technologies,” *INFORMS Transactions on Education*, 2005, 5(2), <http://ite.pubs.informs.org/Vol5No2/Miyaoka/>.
- [7] Serman, J. D. “Teaching takes off: Flight simulators for management education,” *OR/MS Today*, 1992, October, 40–44.
- [8] Wood, S. C. “A survey of publicly available online games for management science, operations, and supply chain management,” *POMS Conference Workshop*, 2006, April.
- [9] Wood, S. C. “Online games to teach operations,” *INFORMS Transactions on Education*, 2007, 8(1), <http://ite.pubs.informs.org/Vol8No1/Wood/>.

USING THE TIME WISE™ SIMULATION TO TEACH LEAN FUNDAMENTALS IN THE INTRODUCTORY OPERATIONS MANAGEMENT COURSE

Susan E. Pariseau, Merrimack College, 315 Turnpike Street, North Andover, MA 01845,
(978) 837-5417, susan.pariseau@merrimack.edu

Sharon Johnson, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609,
(508) 831-5183, sharon@wpi.edu

ABSTRACT

This paper reports results from the use of the Time Wise™ simulation to teach lean principles in the introductory operations management course and compares assessment results for two groups of students, those participating in the simulation and those hearing a traditional lecture. An analysis of the assessment data shows that the simulation group was more likely to include waste and flow in their explanation of lean but less likely to include pull. Whereas both groups, on average, identified the same number of examples of waste, the simulation group appeared more able to apply their learning in a new setting.

INTRODUCTION

There have been numerous articles written describing experiential learning, providing examples of its use, and discussing the pros and cons of this type of learning program. Research in education has clearly demonstrated that active participation by students in the educational process increases learning and retention. In an active and reflective learning environment, students take more responsibility for their own learning. The teacher becomes a facilitator and coach rather than a lecturer and the focus is on guided learning and exploration. The essence of experiential learning and the rewards attributed to it are best described by a quote attributed to the Chinese philosopher Confucius (551 BC – 479 BC):

I hear and I forget. I see and I remember. I do and I understand.

As early as mid-1930s, Dewey advocated an active, student-centered approach to learning (Hickcox, p. 123). Kolb's research found that students learn in different ways (Kolb, 1984) thus reinforcing the idea of using a variety of different teaching techniques in order to accommodate varying learning styles and preferences. Whereas, some students prefer a passive, lecture environment, others prefer to be actively involved in the learning process.

Various authors have reported the results of using some form of experiential learning in their classes. Studies reporting the results of the use of experiential learning in statistics have reported increased retention (Kvam 2000), improved performance on examinations (Hakeem 2001; Pariseau and Kezim 2007) and an enhanced level of satisfaction with the course (Rosenthal 1995). Clements (1995) reported that students in a developmental psychology course that used experiential learning activities found the material more valuable and more interesting. Fink (2007) reports that the use of active learning in a computer engineering course on coding theory increased student motivation dramatically as students "worked harder and enjoyed it more" (p.

27). Hawtrey (2007) found that economics students indicated a preference for experiential learning activities.

At one institution in the northeast region of the United States, the use of experiential learning to teach the fundamentals of lean using a hands-on simulation was first employed in the Operations Management course during the spring 2007 semester. The remainder of this paper will discuss the simulation used and will report on differences found between those students involved in the experiential exercise and a control group.

METHODOLOGY

During the spring 2007 semester, four sections of the introductory Operations Management class were taught. Most of the class meetings were identical in format for all four sections. Only the methodology for the coverage of lean operations varied. Students in two sections participated in the Time Wise™ simulation, an experiential learning format, and two sections were taught lean using a traditional lecture format. Due to availability of the Time Wise™ simulation kit and the length of time required for setup, the two classes taught on Wednesdays participated in the simulation and did not receive a formal lecture on the topic of lean. The two sections taught on Fridays did not participate in any simulation activities but did hear a lecture on lean. It should be noted that the Wednesday and Friday classes were taught at the same time of day (8:00 - 9:45 AM and 10:00 - 11:45 AM). A total of 43 students were enrolled in the sections participating in the simulation; 30 students were enrolled in the non-simulation/lecture sections.

During the spring 2007 semester, knowledge of lean concepts was assessed following completion of the topic. All students completed the same assessment regardless of the teaching style. The assessment required that students view the first segment (8-piece lot size) of the Styro video (provided by John Deere and Company and APICS, distributed by Richard D. Irwin, Inc. 1991). Following the video, students were asked to provide answers to the following questions:

1. Provide a brief explanation of lean principles
2. Provide examples of waste evidenced in the Styro video. If you see the same type of “waste” in multiple places in the video, please list it multiple times and tell where you saw it. Provide as many examples as you can.
3. Explain process changes that you would recommend to make this process more consistent with lean principles. Explain how you would eliminate waste. Be specific.
4. Does Styro use a push system or a pull system? How can you tell?
5. What would you expect to happen if Styro decreased the lot size?

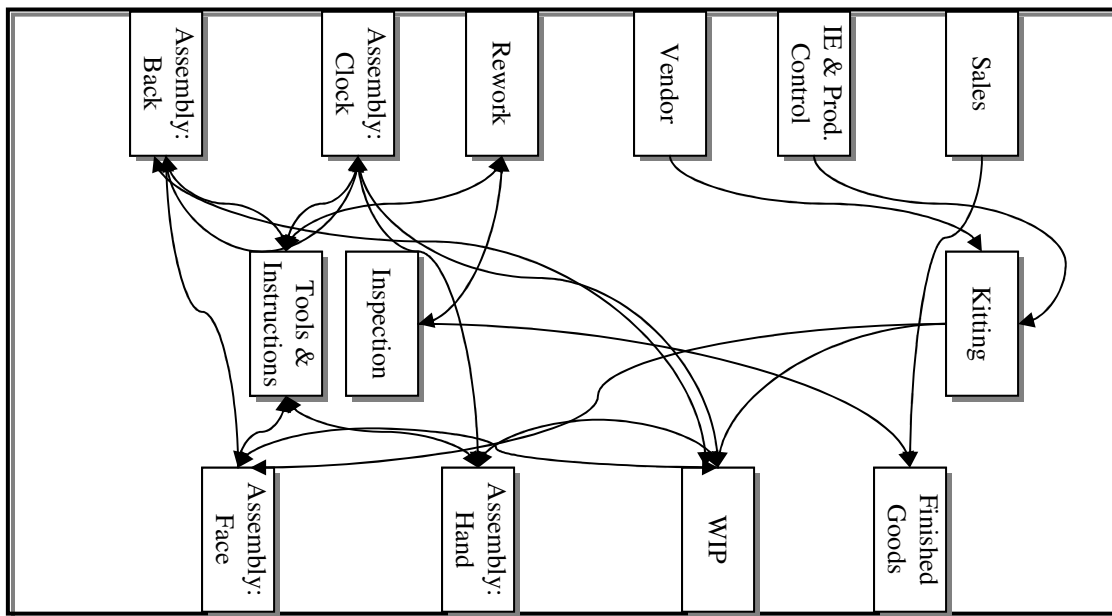
SIMULATION

Through the use of Time Wise™, students participate in a high volume, standard product manufacturing simulation. The simulation, developed by Time Wise Management Systems, has been used extensively in corporate lean training. In the simulation, students assemble simple clocks. During spring 2007, each class section completed four 15-minute simulation rounds, two rounds on each of two days. Between rounds held on the same day, students participated in

group discussions to explore possibilities for process improvement and the instructor made brief presentations to introduce lean concepts.

The initial simulation round provides a scenario with traditional manufacturing processes. Figure 1 provides the layout, flow and positions filled by students in this round. In this first round, the material handler moved all materials from work center to work center. In subsequent rounds, positions were eliminated or modified and the layout was changed according to student suggestions for improvement. Improvements made included balancing the line, changing or redefining jobs, eliminating positions, and decreasing lot sizes.

FIGURE 1. Initial Layout, Flow & Positions



Students collected and analyzed performance data to determine how well their company was meeting their competitive priorities of quality and delivery. In many cases, especially following round 1, the students discovered that they had collected both inaccurate data and an insufficient amount of data. As a team, they were challenged to find methods to improve data accuracy to aid in the identification and elimination of additional waste.

After each new round, additional fundamentals of lean were briefly presented in an effort to guide students to think in new ways. The cycle was one of continuous improvement and the pursuit of perfection and zero waste. It was clearly a reflective cycle that provided students opportunities to work in a team while learning from their experiences. They were required to reflect on what worked and what did not work. As is the norm in experiential learning, the instructor acted as a facilitator and coach and did not “tell” the students exactly what they should do to make improvements.

Typically, the students identified the jumbled flow of the initial layout and improved the flow in Round 2. They also generally eliminated the positions of instruction and tool crib attendant and requested that both tools and instructions be placed at the appropriate work centers. In Round 3,

students eliminated kitting, used point-of-use storage for all parts and attempted to eliminate bottlenecks and balance the line. During Round 4, guided by the instructor, kanban use and pull manufacturing was utilized. By the completion of four rounds, most of the main ideas embraced by lean practitioners were implemented. The process used was one of experimentation and guided learning.

As part of the Time Wise™ simulation kit, a spreadsheet is provided for data input (similar to that shown in Table 1). Financial reports flow from the data input. The data provided below, taken from one of the spring 2007 sections of Operation Management, provides evidence of a decrease in defect rates, a decrease in WIP, and an increase in on-time delivery as students progressed from Round 1 to Round 4. All of this was accomplished with less space and fewer employees, while introducing a third variation of the product. In Round 1 the group experienced a loss of \$645.00 whereas in Round 4, they had a profit of \$1,515.00. There was little variation in total operating costs but there was a dramatic increase in revenue.

TABLE 1. Data Collection from Rounds 1 & 4

	Round 1			Round 4		
	Product 1	Product 2		Product 1	Product 2	Product 3
Average Lead Time (min.)	>15					
# Units in Ending WIP	36	60		5	5	3
# Units on Time	3	4		36	35	21
# Units Shipped Late	4	0		1		
# Units in Finished Goods	4	0		4	4	3
# of Employees	19			17		
# of Tables	13			6		
# Ft. Traveled	134			24		
# Failed	18			0		
# Passed	18			93		

ASSESSMENT

Rubrics were developed to assess student responses to each of the questions posed after they watched the Styro video. The rubric for analyzing the results of question 1 was based on the fundamentals of lean as defined by Womack and Jones (2003). For each of the fundamental principles of lean, the proportion of students who included this principle in their explanation was tabulated. The proportions were compiled for the two groups of students, those participating in the simulation-only and those hearing the lecture-only. Z-tests for the difference in two proportions were completed for each of the five fundamental principles. As shown in Table 2, significant differences were found between the two groups for 3 of the 5 principles. A significantly greater proportion of students in the simulation group included waste and flow in their explanations. This is not surprising since most of the changes effected during the 4 rounds of the simulation exercise were based upon elimination of non-value adding activity or waste and upon an attempt to make the product flow through the system. It is surprising to see that so few students, especially those participating in the simulation, included the idea of perfection or continuous improvement. The simulation required that the students themselves make changes to

the system. The changes suggested by the students resulted in the continuous improvement of their process and all of their performance measures. With respect to the surprising results for pull, where a greater proportion of the lecture group included this principle, this indicates an area where a lecture/demonstration proved more helpful to understanding than the actual use of kanbans during the simulation with only a brief description presented prior to the simulation round.

TABLE 2. Explanation of Lean Principles

Principle	Simulation	Lecture	p-value
Value (customer focus)	28%	27%	0.846
Waste	95%	80%	0.039
Flow	60%	33%	0.023
Pull	33%	60%	0.020
Perfection (continuous improvement)	5%	13%	*

* violated assumption np and $n(1-p) \geq 5$

Question 2 required students to list all examples of waste seen in the video. Students in the simulation group provided an average of 6.2 examples of waste. Students in the lecture group presented an average of 6.1 examples. There was no difference seen between the two groups of students on this question.

Question 3 asked students to recommend changes to make the process seen in the Styro video more consistent with lean principles. Table 3 summarizes student recommendations and the proportion of students from each of the two groups (simulation and lecture) that list the change. Hypotheses of no differences between proportions of students in the simulation and lecture groups were tested using the Z-test for differences in proportions. For nearly all changes recommended, a greater proportion of students in the simulation group recommended the change. It appears that students who have actually experienced “making a product” can more readily identify opportunities for improvement. Statistically significant differences were found with respect to recommendations for layout changes and point-of-use storage. Both of these changes were made between the first and second round of play. Mildly significant differences were found in recommendations for modification of operations and preventive/productive maintenance. These changes were typically made in the third round of play. Just as the lecture group was more likely to include pull in their explanation of lean, they also more often identified that a switch to a pull system would improve performance.

Question 4 seeks to determine whether students understand the difference between push and pull systems. 91% of the simulation group and 77% of the lecture group correctly classified Styro as a push system (p -value = 0.099). In answering question five, 81% of the simulation group and 53% of the lecture group associated a decrease in lot size with a reduction in lead time (p -value = 0.010).

TABLE 3. Changes Recommended to Make Process Lean

Process change	Simulation	Lecture	p-value
Balance capacity	88%	83%	0.538
Layout change (reduce distance, cell)	63%	20%	0.000
Set-up reduction	44%	33%	0.351
Reduce lot sizes	21%	17%	0.649
Modify steps (eliminate, reorder, combine)	56%	33%	0.058
Preventive/productive maintenance	49%	27%	0.057
Point-of-use storage (POUS)	49%	23%	0.027
Product design (stickers replace markers)	14%	0%	*
Pull system (produce to customer demand)	40%	50%	0.375
Quality at the source	12%	10%	*
Poka-yoke	14%	3%	*
Better information flow(communication)	12%	0%	*
Better inventory planning	26%	23%	0.826
Visual signals	2%	0%	*

* violated assumption np and $n(1-p) \geq 5$

LIMITATIONS AND CONCLUSIONS

Anecdotal comments by students in the simulation sections suggest that they enjoyed participating in the simulation and believed they learned far more than they could have learned in a lecture environment. An analysis of the assessment data shows that this group was more likely to include waste and flow in their explanation of lean but less likely to include pull. Whereas both groups, on average, identified the same number of examples of waste, the simulation group was more likely to propose specific changes to improve the process seen in the Styro video. They also more readily identified the relationship between a reduction in lot size and lead time. It is interesting to note that the lecture group was more likely to include pull in both their explanation of lean and in their recommendations for change, yet they were less likely to recognize that Styro was a push system.

It appears that the simulation and experiential learning format was more effective in allowing students to transfer their knowledge and apply their learning in a new setting. It must be noted, however, that the situation portrayed in the Styro video was similar to the first round of the Time Wise™ simulation. Assessment using different scenarios should be used to determine whether this knowledge transfer would occur in other circumstances. In addition, although it is believed that both the simulation and lecture groups entered the course with little knowledge of lean, no pre-test was conducted.

REFERENCES

References are available from Susan E. Pariseau upon request.

ALCOHOL, A MOTIVATING THEME FOR INTRODUCTORY MATHEMATICS AND STATISTICS COURSES

Reza Noubary

Department of Mathematics
Bloomsburg University
Bloomsburg, PA 17815
rnoubary@bloomu.edu

ABSTRACT

This article describes an ongoing effort to improve teaching and learning of introductory mathematics and statistics for non-majors. It uses alcohol as a motivating theme and utilizes data and information related to it for developing lessons. So far a few lessons are developed and used for teaching of certain classical topics. To cover a wider range of topics, several reports are also prepared to help developing further lessons. The article concludes by discussion of two demonstrating examples.

INTRODUCTION

Teaching introductory mathematics and statistics to non-majors has always been a challenge to instructors. Here the challenge lies in how to tie the material to students' backgrounds or interests so that the subject matter becomes relevant and useful. Recent studies show that teaching of the applied quantitative methods in a context familiar or of interest to students increases their desire to learn and enhances their tolerance for dealing with complex concepts. Considering this, in an ongoing project, I have been investigating an idea for motivating students in introductory mathematics and statistics classes. The motivating theme is alcohol and teachings are based on methods involved in studying and analyzing data and information related to it. The plan is to develop series of lessons for teaching topics covered in introductory mathematics and statistics courses. Based on students' responses to the lessons completed so far, I can report that alcohol is an effective motivating theme.

Work Completed

Several lessons are developed and are written in the form of short reports. They are listed below. The list also includes reports that contain information for developing further lessons.

1. Mathematics of Alcohol and Drinking
2. Alcohol - Facts, Statistics, Resources, and Impairment Charts
3. Hydrometer
4. A Simple Method for Calculation of the Blood Alcohol Content

5. Comments on ‘Substance Use Among College Athletes: A Comparison Based on Sport/Team Affiliation’ by Jason A. Ford, PhD
6. College Students and Alcohol Abuse Statistics
7. Teaching Descriptive Statistics Using Alcohol and Youth Facts
8. Teaching Sample Space and Events Using Greater Dallas Alcohol and Drug Abuse Data
9. Teaching Correlation Using Alcohol Consumption and the GPA
10. Teaching Confidence Interval and Test of Hypothesis Using Alcohol Data for the Memorial Day Holiday
11. Teaching Chi-Square Test for Independence Using Alcohol Consumption Data of the Registered American Indians
12. Teaching Prediction of Future Consumption Using Regression
13. Comments on “Alcohol Abuse Initiative”. Sam Houston University, July 2007 Report
14. Binge Drinking – A Continuing Problem on College Campuses
15. Alcohol Statistics in Pennsylvania
16. Alcohol statistics in Bloomsburg University
17. Alcohol, STD, and AIDS
18. Holiday Drinking and Driving
19. Alcohol Controversies
20. Genetic Protection from Alcoholism

DEMONSTRATING EXAMPLES

Mathematics: Information to Develop Lesson on Functions

After a drink, the alcohol begins to enter the bloodstream almost immediately and the blood-alcohol level rises rapidly. Once the person stops drinking, her natural metabolic processes slowly eliminate the alcohol and blood-alcohol level begins to fall. So, how long must one wait after drinking for safe driving?

Experiments show that after drinking 6 beers, 4 glasses of wine or 4 shots of liquor, the blood-alcohol level typically rises to 0.11 gram per deciliter of blood. Thereafter alcohol is eliminated at the rate of 0.02 gram per hour. This means that over time the blood-alcohol level first rises rapidly and then falls linearly. This information is used to develop a lesson to introduce students to concepts such as:

- 1) Functions and their graphs.
- 2) Linear functions.
- 3) Increasing and decreasing functions.
- 4) Maximums and minimums (derivatives)

As a class activity I ask student to find the length of time one needs to wait before driving. This activity requires calculating where the graph of the function describing the blood-alcohol level falls below the legal limit.

To introduce students to multivariable functions I ask them to think about inclusion of factors such as the amount of alcohol consumed, person's weight, etc.

Statistics: Information to Develop Lesson on Hypotheses Testing

Many people think or claim that:

- 1) The moderate use of alcohol (few drinks) does not impair their driving abilities, such as vision, reflexes, and hand-eye coordination.
- 2) Their body weight helps their level of tolerance and sobriety.
- 3) Their female partners' alcohol tolerance is less than theirs.

I have used this information to develop lessons to introduce students to concepts such as hypothesis testing and design of experiments.

A) Hypothesis Testing

Hypothesis 1: Alcohol consumption makes people impaired, regardless of their claims. Moreover, the more a person drinks, the more impaired they become.

Hypothesis 2: As a person's weight increases, the effect of a certain amount of alcohol decreases. That is, the more someone weighs, the less effect one drink will have on them.

Hypothesis 3: A person's sex has no effect on their ability to handle alcohol without becoming impaired.

Testing these hypotheses requires data. To collect relevant data we need to design and implement an experiment. This can be used to introduce students to design and analysis of experiments. The following is an example of what instructor can present.

B) Design of Experiment

Step 1: Select a random sample of individuals who are willing to participate in this experiment. It is better to have the same number of males and females (balanced design) and people of different weights.

Step 2: Test individuals in your sample using, for example, a driving simulation game such as Sega's Outrun and record their scores.

Step 3: Test them again after they had consumed two beers.

Step 4: Test them once more after they have consumed enough beer to raise their blood alcohol level (content) to .06%, one percentage point above assumed legal level to declare somebody impaired.

Note: It takes typical college students approximately four beers to raise their blood alcohol content to 0.06%.

As a class activity I ask student to discuss the possible problems and biases of this experiment. For example, some subjects "video game experts" may score better on video games than others. Some might have played Sega's Outrun before, while others never saw it. I tell them for example that to account for this one could test everybody before and after drinking and use the difference of the scores. I use this to teach concepts such as dependency and the paired experiment. Next I ask student to think about the effects of factors such as individual alcohol tolerances and its dependency on previous exposure to alcohol, weight, and gender.

Notes: We actually performed the experiment described above using a group of friends during a three month period and draw the following conclusions. I used this to introduce t -test , χ^2 test and regression

Hypothesis 1: The single sample t -tests provided sufficient evidence to conclude that 2 beers and 4 beers have a significant effect on an individual's ability to drive. It can also be noted that the test of the effect of 4 beers is more significant than the effect of only 2 beers. This indicated that the more alcohol a person consumes the more impaired the person becomes. I also used regression analysis to determine the effect, if any, of alcohol on an individual.

Hypothesis 2: Unfortunately the sample size of the data was not large enough to yield any significant χ^2 tests. However, the χ^2 value for the effect of 4 beers is significantly larger than the χ^2 value for the effect of 2 beers. This indicates that the number of beers consumed increases the weight of the individual consuming them becomes more significant.

Hypothesis 3: There were two tests performed to analyze the data for this hypothesis. Again, the χ^2 tests were insignificant due to the small sample size. Although the paired t -test were also insignificant they supported our claims that gender has no effect on a person's ability to handle alcohol without being impaired. Note that the participants were divided by gender. This same division could have been achieved if the individuals were divided by weight, since the women were the lightest and the men were the heaviest. To better this experiment a larger group of individuals that were approximately the same weight should have been taken.

Experiential Learning in a Statistics Class Using Racetrack Betting Markets

Jack Rappaport, La Salle University, (215) 991-3555, rappapor@lasalle.edu

Stephen B. Richter, La Salle University, (215) 991-3555, richter@lasalle.edu

ABSTRACT

This paper describes an experiential learning case that uses the horse race betting markets as a vehicle to study some basic statistical concepts in a standard undergraduate statistics class. The experiment consists of using the class as a mini-simulcast center whereby races are simulcast to the class from the Internet and projected onto a big screen. The students also have access to individual terminals where they can access the real time wagering (betting) pools of a particular race and use an Excel spreadsheet to generate statistical results.

Introduction

There have been many attempts by statistics teachers to reduce the boredom in the classroom using innovative teaching techniques. Foster and Stine [6] developed a classroom simulation in which groups of students roll dice to simulate the success of three different investments. The groups that have the greatest success become the “Warren Buffetts” of the class. Schuster [18] uses a two-player game consisting of repeated tosses of a loaded coin until one of the player’s three outcome sequences occurs. Schell [16] developed ways to use baseball to teach statistics. In their book, *Teaching in Statistics: A Bag of Tricks* [8], Gelman and Nolan have taken steps toward inspiring statistics teachers to do more. One of the goals of using interesting applications is to increase the chance that the student will remember the lecture over a long period of time. Pagano [11] reflects on the fact that there were few lectures that he remembers when he was a student, but he does remember a physics professor swallowing some helium and then speaking to the class in a squeaky voice. He notes that teachers of statistics can be at a disadvantage in that statistics is not a bench science and they do not have readily made experiments that they can perform in front of the class.

In this paper we describe an experiment that uses the horse racing betting markets as a way to teach the students some basic concepts of probability and statistics. This certainly would qualify as an interesting application for students to learn some fundamental concepts of probability and statistics. Many students may never have been to a racetrack or watched a live race. We use an experiential learning approach whereby we use the class as a mini-simulcast center whereby races are simulcast from the Internet. The students would also have access to individual terminals where they can access the real time wagering pools and use Excel spreadsheets to generate statistical results.

Experiential Learning Techniques

The use of experiential learning has been fairly widespread not only in traditional management courses but also in statistics classes. The use of experiential learning techniques has been shown to be important to the teaching of statistical concepts [9] [25]. Hakkem [9] argues that introductory statistics courses need to place greater emphasis on practical understanding by the use of exercises involving real data in order to motivate students and to develop the required skills of analysis and inference. Active learning experiences can provide students with valuable opportunities to apply theory and practice. Traditional modes of delivery in colleges and universities encourage relatively passive styles of learning. Students attend lectures and take notes, and are not encouraged to engage with the material [5]. Such passive learning techniques have been shown to be limited because they do not encourage students to process information actively [15]. This may be particularly problematic in the teaching of statistics [1] [9] [25]. A secondary problem associated with the teaching of statistics is that statistics courses generally result in high levels of anxiety among students, especially those that are not especially mathematically or computer-oriented. Any techniques that can be used to relieve this anxiety would be beneficial to these students and could benefit the overall learning process.

Requiring students to apply theory to real-life situations can engage them in higher order thinking and may encourage long-term memory retention as they personalize the subject

matter [2]. We believe that the experiment that we propose has these components. Although not every student may be interested in horse racing, our experience shows that the novelty of this application can elicit a large interest in the subject. Once the students are engaged in the process of understanding the concept of probability and odds, making choices, and experiencing the results or outcomes, the process can become very personalized and hence can have a larger long term impact in terms of memory recall. Almost every experienced horseplayer remembers the first time he or she visited the racetrack. In our context, we can incorporate this first-time experience with the educational process to teach the basic notions of probability theory and statistics.

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 [3]. The personal selection of horses and the watching of the live race can certainly be considered a “human moment”, and the use of teams in our experiment allows the students to engage in a sense of community. A field trip to the racetrack or the off-track betting center could be considered a full extension of this concept (our university does sponsor a so-called “city in the classroom program” whereby academic learning is supposedly enhanced by taking the appropriate field trips).

Many researchers have noted that student learning can be increased by the use of multi-media technology. Velleman and Moore [24] examine the usefulness of multi-media technology for teaching statistics, with attention to both promises and pitfalls. They indicate that the use of multi-media technology (a system that has at least several aspects of text, sound, still images, full-motion video, animation, and computer graphics) has considerable promise for teaching statistics. Pan [13] predicts that the most effective way to make students learn better is to increase the use of computer technology.

In general, intellectual development has been shown to be largely achieved through active engagement and interaction with others [7] [10] [12]. The term "blended learning" is used to refer to a range of integrated learning activities and is used in diverse ways by different people. The most common understanding of the term seems to be that blended learning enhances different delivery modes and, in particular, on-line and face-to-face learning. Blended learning can also refer to the teaching methods employed, the learning experiences being designed for the students, or the locations from which the learning events take place. In our application, the races are used as application of the theory of probability and statistics and are simulcast into the classroom via the use of the Internet. In principle students could be taken on a field trip to the racetrack and experience the "live" action. The use of the Internet technology essentially reduces the cost of such a field trip and plays the same role that traditional distance learning techniques have been used to reduce the cost of higher education (i.e. the notion of producing a quality education with fewer resources). In our application we add an additional layer of insight into the delivery of the education, when we consider a visit to the live racetrack (or off-track betting site) and a process of engagement with the actual participants (or gamblers) of the sport.

Academic Research

There has been a fairly significant amount of research in finance journals that have carried out statistical studies on the horse betting markets. This adds a sense of legitimacy to using

this application in a teaching experiment in an academic environment. Many researchers have found that the horse race betting markets provide a unique environment to study risk preferences of investors; virtually all of the empirical work analyzing risk preferences in the horse betting markets show the representative bettor to be risk loving [19] [22]. This has resulted in what is called the ‘favorite-long shot’ bias, an established feature of betting markets, whereby long shots win less often than the subjective probabilities imply and the favorites win more often. Bettors tend to prefer long shots to favorites because they prefer the possibility of getting a large return. There has been a considerable amount of research studying this phenomenon [4] [14] [17] [20] [21] [23].

The following information from Rhoda, Olson, and Rappaport [14] illustrates this basic concept. The overall mean return and variance of betting on horses of different odds groups have been calculated based upon a sample of 8,021 horses involving 939 races at Philadelphia Park and Garden State Park over a twelve year period.

Odds Grouping (mid-point)	Number of Horses	Mean Return (per dollar wager)	Variance
.25	6	-.033	.467
1.00	268	-.096	1.107
2.00	583	-1.70	1.807
3.25	1134	-.105	2.958
5.00	1179	-.237	3.909
8.00	1513	-.202	6.280
15.00	1632	-.200	11.009
40.00	1706	-.220	27.016

Note that the mean return per dollar bet is on the average higher for the lower odds horses. (also note that the mean return per dollar is negative because of the track take of about 18% which is deducted from the pool for expenses, taxes, etc.)

Field Research

The experiment has some characteristics of a field trip; even though the students are not physically at the racetrack or the off-track betting site, the Internet technology provides an atmosphere of a real life-betting environment. The experiential learning model could be

viewed as an extension of the day to day experiences that real players have when they gamble at the track. Since both of the authors are frequent players at the track, we can describe some of the aspects of probabilistic events that players can experience at the track on a day to day basis.

One of the classic applications of probability theory is the notion of the binomial distribution. The distribution is based upon having two possible outcomes and then calculating the probability of obtaining different outcomes in a series of independent trials. For example one could calculate the probability that there would be at least 6 heads in a tossing of a fair coin 10 times. In horse racing, the favorite wins the race approximately 30% of time and this statistic seems to have held up remarkably well over the years under all kinds of conditions and tracks. Hence, in a typical card of about 9 or 10 races, one would expect an average of about three of the races to be won by the favorite. There is always a probability (although very small) that not one race on the card is won by the favorite and an even smaller probability that every race will be won by the favorite (essentially zero probability, but theoretically some chance). Players that tend to bet longshots will tend to boast about their preferences when they see a streak of favorites lose; and players that tend to bet favorites will often wager larger amounts of money on the favorite after a streak of them loses, hoping that the law of averages will turn in their favor. But the law of the binomial distribution states that one cannot predict the probability of the next trial based upon what happens in the preceding trials, due to the law of independence, (each horse race could be considered independent of each other). Of course the law of the binomial probability distribution could also be applied to other forms of gambling such as roulette or dice games.

Another interesting phenomenon is to observe player behavior in relation to the so-called 'favorite-longshot' bias established in the literature. As explained above research has shown that the representative bettor has been shown to be risk loving, i.e. he is willing to accept a lower expected return and higher variance for the opportunity, albeit with low probability, to earn a significant payoff. In simple language this means that the typical horse player tends to prefer larger odds choices because he gains utility from the hope of a large payoff. A casual observation and engagement with the typical players at the track will confirm this behavior. Many players tend to view playing the favorite as a "suckers" bet and will often chastise another player who wagers a considerable amount of money on the favorite. They will say something like, "you are such a fool to bet so much on the chalk (i.e. the favorite). Moreover, I have seen so many of them lose and even if you win you aren't going to get anything for your money".

The reality is actually just the opposite. As pointed out above, statistical studies have shown that the mean return per dollar bet is actual higher betting the lower odds horses than betting the larger odds horses. It could very frustrating for a statistician to explain to a typical horseplayer why he would be better off in the long run betting on the lower odd horses. The racetrack takes advantage of this bettor behavior as they structure the game. The exotic wagers generally yield very large payoffs and are based upon selecting the various combinations of horses coming in the correct order. For example an "Exacta" bet consists of selecting the first two finishers in the race in the correct order; a "Pick Three" involves selecting the correct winner of three races in a row. The exotic wagers generally attract more money than the straight win, place and show wagers. Bettors are attracted to these wagers because of the potential of a large payoff. The track generally takes a larger

percentage of the wagering pool on these kinds of wagers. So in essence the mean return of these bet is lower over a long period of time compared to the standard win, place and show wagers.

On the flip side, some bettors who wager large sums of money tend to prefer favorites particularly if their subjective estimate of the probability that the horse is greater than the subjective probability of the overall market (i.e. the one implied by the odds). A situation that occurs at the track on a regular basis is when an individual wagers an exceedingly large amount of money on what he considered a “sure” thing. For example it is not uncommon to see huge bets made on the show pool on a horse that seemingly must come in at first, second or third; players that wager in this fashion are called “bridge jumpers”. Although they choose low odds, high probability choices, they also are risk takers because the amount of money that they wager can be considered to large compared to their total bankroll. As every gambler knows, there is no such thing as a sure thing.

An Experiential Approach to Probability Theory

The concept of probability may seem intuitive obvious to most statistics teachers but may elude some students, particularly those taking statistics for the first time. A probability is the numeric value representing the chance, likelihood, or possibility that a particular event will occur. The most direct explanation of probability is to refer to it as the percentage of the time that the event would occur under the same circumstances. For example the probability that it will rain tomorrow is .4 is the same thing as saying that in the long run, 40% of the time it will rain the following day under exactly the same circumstances that exist today.

An experiential approach to the study of probability using the horse race betting market can be potentially very useful in helping students better understand this concept. A process of active participation of selecting horses with given odds (or probability) can allow the students to apply the theoretical concept in a real-life application. A traditional approach to the study of probability theory considers the difference between “a-priori classical probability” (such as the probability of obtaining a head on the toss of a coin), versus “empirical probability” where the outcomes are based upon observed data (such as the proportion of households that have purchased or plan to purchase a big screen television) and “subjective probability” which is based upon a combination of an individual’s past experience, personal opinion and the analysis of the particular situation and can therefore differ from person to person.

In our horse racing experiment, the probability that a given horse wins the race could be considered a “subjective probability” and could be computed by the percentage of the money that is bet on that horse out of the total wager pool. The wagering pools (at least for win, place and shows bets) are generally available on the simulcast programs and thus can be viewed on the Internet in real-time environments. The notion of “empirical probability” can also be illustrated since information can be collected from a large sample of races to estimate the probability of certain outcomes. For example it is well known among horse players that the probability that the favorite (i.e. the horse with the lowest odds as displayed on the tote board) wins the race approximately 30% of the time and this statistic seems to have held up remarkably well over most tracks over a long period of time.

The notion of “subjective probability” can be studied in more depth when one considers the odds based upon the overall betting pool, the odds that are assigned by the morning line handicapper, and an individual’s own personal assessment. The morning line odds represent the best guess by an expert in the field as to what the final odds will become, whereas the actual “final odds” represents the best “market efficient” estimate of the probability since it involves the opinion of the entire spectrum of bettors. An individual may decide to choose a particular horse if he or she believes that the probability that it wins the race is greater than the probability projected from the betting pools. The notion of “efficient markets” would indicate that one couldn’t improve upon the assessment of the entire market unless one has inside or special information. Of course, the normal handicapping information could be made available to the students. This consists primarily of the past performance of the horse and some relevant breeding information. Part of the fun of horse racing is for the spectators to make their own personal choices based upon their analysis of the past performance information.

An interesting experiential aspect of horse race betting involves the analysis of the dynamic process by which the odds change over the betting period (usually a 20-minute time block). Rhoda, Olson and Rappaport [14] studied the odds changes over the course of the betting period and found that the information flows from the tote board could be used to earn abnormal and possibly positive returns. Many racetrack bettors actively watch the betting patterns on the board and look for insights for their betting choices. It is possible that more subtle changes have more informational content than large ones. For example, a significant drop in the odds at the very end of the betting period may cause a bandwagon effect to bet on that horse. (The “efficient markets” hypothesis would indicate that the bettor couldn’t improve his long term return by following such events).

Another relationship often followed by the bettors is that between the actual “real-time” odds and the morning line odds. For example a horse whose actual odds are significantly below the morning line odds may attract some attention from the bettors since they believe that there are bettors with special or inside information betting on this selection. Other bettors might reach an entirely different conclusion and believe that the odds are too low and that their estimate of the probability of the horse winning is lower than what is implied by the tote board.

Another concept in probability theory is to distinguish between a “discrete” and “continuous” random variable. Students can easily see that the selection of the winning horse involves the application of discrete random variables since we are dealing with outcomes that result from a counting process (i.e. which horse will win the race).

All of these examples show how rich the study of horse racing can be for first time students in a probability and statistics course. An experiential involvement in the betting process can give the student an “active” understanding of the notion of probability and help them gain a better theoretical understanding of some of the concepts.

Now we will describe the Excel spreadsheet programs that will be used by the students to carry out various computational formulas. Using the Excel spreadsheets during the experiment can also be considered a form of experiential learning. Students would be expected to enter data using the spreadsheets to record the results of the races that they selected. The statistics generated by the spreadsheets would change as the betting process evolves. The students could also use the spreadsheets for “what-if” analysis.

Probability Distributions

The following spreadsheet can be used to understand the concept of a probability distribution. The wagering pools for win, place and show bets can be entered into this spreadsheet and the spreadsheet calculates the percentage of the win pool bet on each horse as well as the final odds that would result from the money wagered. The percentage of the money bet on each horse could be interpreted as the subjective probability that the horse wins the race. This would represent the probability distribution of the various outcomes of the race. In a probability distribution, the sum of the probabilities of all possible outcomes is unity or 1. This is very easy to understand in this example since the total pool is the sum of all the money wagered on each individual horse and the probability that each horse wins is the percentage of the money wagered on that particular horse.

The spreadsheet can also be used to carry out sensitivity analysis. For example one could determine how the probabilities or odds change if the amount wagered on a particular horse were to increase by a certain amount; one could also find the impact that changes in the track take would have on the final odds.

BELMONT PARK - 09-07-07		DATA: COURTESY OF www.tvgfree.net				
RACE # 1 HORSE NUMBER	HORSE NAME	WIN POOL	PLACE POOL	SHOW POOL	WIN %	FINAL ODDS
1	CHAMELEON	\$4,675	\$2,612	\$1,474	3.0912%	25.53
2	RIVER MOUNTAIN RD	\$4,432	\$2,206	\$1,537	2.9305%	26.98
3	LETHIMTHINKHESBOSS	\$13,401	\$6,135	\$3,068	8.8609%	8.25
4	TACTICAL GOLD	\$70,188	\$20,142	\$10,762	46.4093%	0.77
5	BAILERO (ARG)	\$32,325	\$9,786	\$4,559	21.3737%	2.84
6	KEE KAW	\$8,000	\$3,386	\$1,763	5.2897%	14.50
7	WOOD WINNER	\$5,813	\$2,537	\$1,263	3.8436%	20.33
8	HALO'S TIGER	\$3,504	\$1,557	\$768	2.3169%	34.39
9	FIT FAZE	\$8,899	\$3,324	\$1,963	5.8841%	12.94
TOTALS		\$151,237	\$51,685	\$27,157	100.0000%	

Track
Takeout

0.180

RESULTS

	7-Wood Winner	\$44.20	\$11.40	\$7.60
2nd	4-Tactical Gold	-	\$2.90	\$2.40
3rd	6-Kee Kaw	-	-	\$5.90
4th	5-Bailero (Arg)	-	-	-
\$2.00 EXACTA 7-4 \$122.00				
\$2.00 TRIFECTA 7-4-6 \$997.00				

Expected Value and Standard Deviation

The following spreadsheet would be used by the students to calculate the mean return per dollar bet and the standard deviation per dollar bet for various wagering possibilities. Students would enter the name of the horse, the amount bet, the final odds, and a 0 or 1 depending upon whether the horse lost (0) or won (1) the race. The program would calculate the total amount wagered, the amount won, the expected return per dollar bet, as well as the standard deviation per dollar bet.

INPUT TABLE						
HORSE NAME	FINAL ODDS	AMOUNT WAGERED	WIN/LOSS (0 OR 1)	RETURN PROFIT	RETURN PER \$ BET	SUM OF SQUARES
CHAMELEON	2.50	\$10.00	1	\$25.00	\$2.50	\$5,958.04
RIVER MOUNTAIN	5.00	\$2.00	0	-\$2.00	-\$1.00	\$13.43
LETHIMTHINK	1.20	\$2.00	0	-\$2.00	-\$1.00	\$13.43
TACTICAL GOLD	3.50	\$2.00	0	-\$2.00	-\$1.00	\$13.43
BAILERO (ARG)	2.80	\$2.00	0	-\$2.00	-\$1.00	\$13.43
KEE KAW	3.20	\$2.00	0	-\$2.00	-\$1.00	\$13.43
FIT FAZE	5.50	\$2.00	0	-\$2.00	-\$1.00	\$13.43
		\$22.00		\$13.00		\$6,038.59

OUTPUT TABLE

TOTAL PROFIT	TOTAL AMOUNT WAGERED	MEAN RETURN (per \$ bet)	STANDARD DEVIATION (per \$ bet)
\$13.00	\$22.00	\$0.59	\$16.96

The use of this program would expose the students to an experiential approach to understand the concept of expected value and standard deviation. The concept of expected value can be illustrated by calculating the long-term expected return by betting on horses in different odds groups. The historical data above estimates the overall mean return and standard deviation for different odds groups calculated from a sample of 8,021 horses. These numbers can be used as benchmarks for the long-term average and standard deviation for the return on betting on the different odds groups. During the experiment the students could then compare the mean return and standard deviation for their particular sample with these values. (For the initial experiment students could use the same bet size for each horse in order to make these comparisons).

The students could perform a sensitivity analysis of the effect on the overall return by adding or deleting a particular horse from the sample. The program also allows for the entry of different bet sizes for the different selections. This would allow the students to perform a sensitivity analysis on the effect on the mean return and the standard deviation of return by increasing the bet size on particular horses. Real horse race gamblers

experience this phenomenon on a frequent basis and often do not realize the extent that they are increasing the variation in the return by increasing the bet size on particular selections.

The Sampling Distribution

The next spreadsheet calculates the 99% limits for the sampling distribution for horses with different odds groups and sample sizes. A sampling distribution is the distribution of the mean return of you selected all possible samples. For example one could determine the distribution of returns based upon wagering on a sample of 50 horses.

99% LIMITS FOR MEAN RETURN			SAMPLE SIZES	SAMPLE SIZES	SAMPLE SIZES	YOUR SAMPLE SIZE
ODDS GROUPS	MEAN RETURN	VARIANCE	50	100	500	1000
1.00	-0.13	1.11	-0.58	-0.45	-0.27	-0.23
			0.32	0.19	0.01	-0.03
2.00	-0.13	1.80	-0.70	-0.53	-0.31	-0.26
			0.44	0.27	0.05	0.00
5.00	-0.21	3.90	-1.05	-0.80	-0.47	-0.40
			0.63	0.38	0.05	-0.02
8.00	-0.21	6.30	-1.27	-0.96	-0.55	-0.45
			0.85	0.54	0.13	0.03
15.00	-0.21	11.00	-1.62	-1.20	-0.65	-0.52
			1.20	0.78	0.23	0.10
40.00	-0.21	27.00	-2.41	-1.77	-0.91	-0.70
			1.99	1.35	0.49	0.28

We use the data from [14] to calculate these limits. First we smooth out the results by finding the mean return per dollar bet for the odds groups 3.25 and below and the odds groups above 3.25. This results in a mean return of -.13 for the odds groups below 3.25 and a mean return of -.21 for the odds groups above 3.25. Using the “central limit theorem” we can then use the normal distribution to estimate the 99% limits for the mean return. It is understood that the results would only be approximate for small sample sizes, since the underlying distribution is not normal. The following spreadsheet calculates these 99% limits for the mean return for various sample sizes. For example the results show that if one consistently bet horses in the 2-1 odds category, a random selection of 100 of these bets would result in a 99% probability that the mean return per dollar bet to be between -. 53 and .27.

The concept of the sampling distribution can be considered a fairly sophisticated one for a first time statistics student. This experiment allows the student to participate in an experiential way with the concept of the sampling distribution; by interacting with the outcomes of the races and the spreadsheet program, he gets a sense of the variation of the

sample mean for a selected sample of races. The 99% limits give him a range within which he can expect the sample mean to fall.

Two fundamental principles of sampling theory can be understood by participating in this experiment. The first one is that the larger the sample size the more likely the sample mean is close to the true mean. Another principal is that the larger the standard deviation of the underlying distribution the higher the variation in the sample mean, and therefore the less likely the sample mean will be close to the true mean. Both of these principles can be understood by participating in this experiment.

Experiential Learning Case

Now we describe the full experiential learning case that we have developed using the concepts described above. The experiment consists of using the class as a mini-simulcast center, whereby races are simulcast to the class from the Internet and projected onto a computer screen. The experiment is designed to take place in a computer lab whereby each student would have his or her own computer so that he can individually check the betting pools for each individual race as well as have access to an Excel spreadsheet where he can enter the results of each race and generate the various statistical results described above.

The experiment will use the "Phonebet Simulcast and Wagering System" to simulcast races from around the country. Typically four or five racetracks are simulcast for live races, but the real time betting pools are given for most of the racetracks that are running on any particular day. On a typical day during the week, at least 10 or 15 tracks could be running during the day and another 10 or 15 at night. (Many of the night tracks would be harness tracks but the same principle applies). On a typical Friday or Saturday evening more tracks are available. To make the experiment entertaining and attractive one could restrict the experiment to races where only the live race was shown. Of course it would be very easy to generalize the experiment to include all the races where the real time betting pools can be accessed even though the live race is not shown.

Students would be arranged in teams and asked to make mock bets on the races. The morning line odds, and the real time odds changes are provided for each race. Students can check at any time what the betting pools are for any race running at the time. Teams would be arranged by odds groupings, to represent the different levels of risk for each odds group. For example the first team would restrict their selections to the lower odds group (under 5), a second team would choose middle odds horses (5-15), whereas a third team would restrict themselves to horses with highest odds (over 15). Another team would be given no restrictions and therefore would be free to select horses of any odds category. Although there probably would only be time to include about 10-12 races that could be seen live, multiple wagers could be made on each race, so that it would be possible for 20-30 selections to be made over the course of the experiment. The experiment could be generalized so that literally hundreds of selections could be made if one includes all of the current races as well as the past races.

Each team would enter the results of their selections into an Excel spreadsheet which would then be used to calculate the mean and standard deviation of the return both in total dollars and in per dollar bet.

The following spreadsheet summarizes the results of one of teams that participated in the experiment. This team made selections on a random basis from a variety of races across all the odds groups.

FINAL ODDS	AMOUNT WAGERED	WIN/LOSS (0 OR 1)	RETURN PROFIT	RETURN PER \$ BET	SUM OF SQUARES
2.5	\$10.00	1	\$25.00	\$2.50	\$6,385.38
5.0	\$2.00	1	\$10.00	\$5.00	\$210.92
1.2	\$10.00	0	-\$10.00	-\$1.00	\$946.86
3.5	\$2.00	0	-\$2.00	-\$1.00	\$5.99
2.8	\$2.00	0	-\$2.00	-\$1.00	\$5.99
3.2	\$10.00	1	\$32.00	\$3.20	\$10,413.09
5.5	\$2.00	0	-\$2.00	-\$1.00	\$5.99
3.0	\$25.00	0	-\$25.00	-\$1.00	\$15,290.17
7.0	\$2.00	0	-\$2.00	-\$1.00	\$5.99
8.3	\$3.00	0	-\$3.00	-\$1.00	\$22.37
4.6	\$4.00	1	\$18.40	\$4.60	\$1,394.17
12.8	\$2.00	0	-\$2.00	-\$1.00	\$5.99
19.0	\$2.00	0	-\$2.00	-\$1.00	\$5.99
22.3	\$15.00	0	-\$15.00	-\$1.00	\$3,254.90
18.5	\$6.00	0	-\$6.00	-\$1.00	\$197.04
9.1	\$3.00	0	-\$3.00	-\$1.00	\$22.37
3.2	\$2.00	1	\$6.40	\$3.20	\$88.96
4.6	\$5.00	0	-\$5.00	-\$1.00	\$111.90
7.5	\$8.00	0	-\$8.00	-\$1.00	\$478.11
13.3	\$11.00	0	-\$11.00	-\$1.00	\$1,266.62
12.4	\$14.00	0	-\$14.00	-\$1.00	\$2,639.45
6.8	\$22.00	0	-\$22.00	-\$1.00	\$10,388.90
4.9	\$16.00	0	-\$16.00	-\$1.00	\$3,959.27
5.1	\$3.00	1	\$15.30	\$5.10	\$727.21
8.6	\$8.00	0	-\$8.00	-\$1.00	\$478.11
	\$189.00		-\$50.90		\$58,311.75

OUTPUT TABLE

TOTAL PROFIT	TOTAL AMOUNT WAGERED	MEAN RETURN (per \$ bet)	STANDARD DEVIATION (per \$ bet)
-\$50.9	\$189.00	-\$0.27	\$17.61

Using the spreadsheet one could perform sensitivity analysis to determine how changing the bet size would have affected the overall mean return and standard deviation of return. For example one could determine the effect on the overall return if one of the losing horses had won, or vice-versa. One could also perform a sensitivity analysis on the effect on the overall return if the bet size were changed on a specific horse.

Conclusions

The experiment seemed to work successfully in engaging the students in an active process of viewing the races, recording the data, and generating the statistical results. The most

revealing aspect of the experiment was to observe the excitement that the students showed in watching the races and participating in the experiment. It was clear that many of the students had never been to a racetrack or watched a live race. It would be difficult to measure the exact relationship between this involvement and the potential to learn, but it is very plausible that the experiment would increase the probability that the student would retain whatever information he or she gained over a longer period of time.

The excitement of the process was enhanced because of the large variation in outcomes, particularly for those choosing longshots. The active engagement in a process with such a large variation in outcomes not only creates an exciting and entertaining atmosphere, but also provides a learning environment within which the students can better understand the notion of variance and standard deviation.

The use of multi-media was definitely a very positive aspect of the experiment. The viewing of the races involves sound, full-motion video and in some cases animation. The experiment combines this with the use of a statistical spreadsheet package that potentially could also be incorporated with computer graphics. Thus the experiment engages the students in a full range of multi-media technologies.

The experiment has the potential to increase the students understanding of risk and return in traditional business applications. Many business decisions that are made in an uncertain environment could be viewed much as a horse race. Investment decisions, the launching of a new venture or new product, the building of a facility, etc. are often one time decisions that have either a positive or negative outcome. Although business decision-makers are generally considered risk averse, they still take risks in many of their decisions. So the notion of subjective probability, variation of outcomes and risk can apply to many traditional business problems. Many decision-makers in traditional business decisions may engage a similar experience that racetrack bettors have with regard to the notion of risk and probability theory.

Although business decision-makers are often viewed as risk averse, there may be some that are risk loving and like racetrack bettors enjoy the thrill of choosing alternatives that have a high potential outcome and a low probability of success. The notion of how many resources are allocated to a risky business venture is analogous to the bet size in the racetrack application. And as the example of the “bridge jumper” shows, allocating an extraordinarily large amount of resources to a fairly safe venture still involves a lot of risk. There is no such thing as a “sure” thing, either at the racetrack or in the traditional business world.

References

- [1] Akram,M., Siddiqui,A.J., and Yaseem,F., “Learning Statistical Concepts”, **International Journal of Mathematical Education in Science and Technology**, 2004,Vol 35 No 1
- [2] Anderson,E., “Active learning in the lecture hall”, **Journal of College Science Teaching**, 1997, Vol 26 May
- [3] Burgon H. and Williams D., “Bringing Off-campus Students on Campus: An Evaluation of a Blended Course,” **The Quarterly Review of Distance Education**, 2003, Vol 4 No 3
- [4]Coleman, C., “New light on the longshot bias”, **Applied Economics**”, 2004, London, Mar Vol 36 Iss 4
- [5]Dabbour,K., “Applying active learning methods to the design of library instructions for a freshman seminar”, **College and Research Libraries**, 2001, vol 58 May
- [6]Foster, D.P. and Stine,R.A.,” Being Warren Buffett: A Classroom Simulation of Risk and Wealth When Investing in the Stock Market”, **the American Statistician**, 2006, Feb, Vol 60, Iss 1
- [7]Garrison,D. and Anderson,T., E-Learning in the 21st Century: a framework for research and practice, 2003, Rouldege Palmer, London
- [8]Gelman,A. and Nolan,D., Teaching Statistics: A Bag of Tricks, 2002, Oxford University Press
- [9]Hakeem,S., “Effect of Experiential learning in Business Statistics”, **Journal of Education for Business**, Nov-Dec 2001
- [10]Laurillard,D., ”New Technologies and the curriculum”, in Scott, P.(ed) **Higher Education re-formed**, 2000, Palmer Press, London
- [11] Pagano,M., “StatConcepts: A Visual Tour of Statistical Ideas”, **The American Statistician**, Alexandria, 2000, Nov Vol 54 Iss 4
- [12] Palloff,R. and Pratt,K., 1999, Building Learning Communities in Cyberspace, 1999, Jossey-Bass, San Fransisco
- [13] Pan,W., “The Challenges of Teaching Statistics in the Current Technology Environment”, **Journal of the American Academy of Business**, 2003, Cambridge, Hollywood, Sep vol 3

- [14] Rhoda,K.L.,Olsen, G.L., Rappaport,J.M., “Risk Preferences and Information Flows in Racetrack Betting Markets”, 1999, The Journal of Financial Research, Vol XXII, No 3, Fall
- [15] Sharkarian,D., “Beyond lecture: Active learning strategies that work”, Journal of Physical Education, Recreation and Dance, 1995, Vol 66, May-June
- [16] Schell,M.J.,”Teaching Statistics Using Baseball”, Mathematical Intelligencer, Heidelberg, Winter 2006, Vol 28 Iss 1
- [17]Shin,H.S., “ Prices of State Contingent Claims with Insider Traders, and the Favorite-Longshot Bias”, The Economic Journal, London, Mar, 1992, Vol 102, Iss 411
- [18] Shuster,J.J., “Using a Two-Player Coin Game Paradox in the Classroom”, The American Statistician, Alexandria, Feb 2006, Vol 60 Iss 1
- [19] Snyder,W., “Horse Racing: Testing the Efficient Market Hypothesis”, 1978, Journal of Finance, 33
- [20] Sobel,R.S. and Raines,S.T., “An examination of the empirical derivatives of the favorite-longshot bias in racetrack betting”, Applied Economics, London, 2003, mar 4 Vol 35 Iss 4
- [21] Swindler,S. and Shaw,R., “Racetrack wagering and the “uninformed” bettor: A study of market efficiency”, Quarterly Review of Economics and Finance, Greenwich Fall 1995 Vol 35 Iss 3
- [22] Weitzman,M., “Utility Analysis and Group Behavior: An Empirical Study”, Journal of Political Economy, 1965, 73
- [23] Williams, L.V. and Paton,D.,”Why is there a favorite-longshot bias in British racetrack betting markets?”, TheEconomic Journal, London, 1997, Jan Vol 107 Iss 440
- [24] Velleman,P.F, and Moore,D.S., “Multimedia for Teaching Statistics:Promises and Pitfalls”, The American Statistician,Alexandria, 1996, Aug Vol 50 Iss 3
- [25] Zeis,C.,Shah,A., Regassa,H. and Ahmadian,A., “Statistical Components of an Undergraduate Business Degree: Putting the Cart Before the Horse”, Journal of Education for Business, 2001, Nov-Dec

IT IS TIME TO INCLUDE EXTREMES IN STATISTICS CURRICULUM

Reza Noubary

Department of Mathematics
Bloomsburg University
Bloomsburg, PA 17815
rnoubary@bloomu.edu

ABSTRACT

This article aims to emphasize the need for inclusion of statistical theories of extreme values in introductory statistics courses. It points out situations where extremes values are of greater concern or importance than averages. The theory of exceedances together with a brief account of extreme value theory, threshold theory, and the theory of records are discussed. It is hoped the instructors of statistics find this article of some value for teaching topics related to non-typical and non-average values.

INTRODUCTION

Traditionally, statistics has focused on the study of values with high frequencies and on averages. This is evident from examination of course descriptions and textbooks for introductory statistics courses offered in universities and colleges. In the modern world, however, it is becoming increasingly impractical to focus on averages alone; it has become important to pay attention to rare events and extremes, since they are newsmakers and are often accompanied by severe consequences.

The celebrated *central limit theorem* has given statistics its focus on averages – for we do what we know how to do. The statistical theories of extreme values are less simple, less unified, and more recent. However they are not less important. In fact, there is a need for inclusion of extreme value statistics in introductory courses.

Why Averages?

As pointed out most textbooks focus on values with high frequencies and on averages. This is because most classical approaches treat the data as a message and seek to decompose it into a systematic part and a random part. Another representation is;

$$\text{Message} = \text{Signal (systematic part)} + \text{Noise (random part)}$$

For data with time index, known as times series, the Wald decomposition theorem states the following:

Time Series = Deterministic (Trend) + Stochastic (Noise)

When the form of a signal is known or is assumed, it is easy to separate the two parts. In the absence of information concerning the systematic or random parts, smoothing is used to recover each part. Smoothing is an exploratory operation, a means of gaining insight into the nature of data without precisely-formulated models or hypotheses. This is usually carried out assuming that the systematic part is smooth and the random part is rough, that is

Message = Smooth part (trend) + Rough part (noise)

Smoothing is often achieved by some sort of averaging (low-pass filtering). Several techniques (low-pass filters) are available for doing this. Once the smooth part is determined, the difference between the original message and the smooth part is used to provide an estimate for the rough part. The rough part is usually used to make reliability statement regarding the systematic part.

For time series, for example, one popular smoothing technique is known as moving averages. The idea is to average the neighboring values and to move it. To clarify, suppose that $Y_t, t = 1, 2, \dots, n$ is an observed time series of length n . The moving average of order $2p+1$ is defined as

$$S_t = \sum_{j=-p}^p w_j Y_{t-j}, \quad t = p+1, \dots, n-p$$

where $w_j > 0$ are weights such that $\sum_{j=-p}^p w_j = 1$. The weights are usually selected based on prior information. One classical approach, referred to as exponential smoothing chooses the weights

according to a geometric progression. It is useful when assigning bigger weight to more recent values make sense.

Why Extreme Values?

In many applications, it is not appropriate to focus on averages alone. In fact, there are many instances where extreme values and values with low frequencies are of more concern than average values or values with high frequencies. Examples include; a large natural disaster compared to a moderate or an average one; a weak component or link compared to their average counterpart; a large insurance claim compared to an average claim. In sports, of course, fans only remember out of ordinary or exceptional performances or games.

So, one may ask why then extremes are not yet found their place in classical textbooks? There are several reasons for this; maybe because the celebrated *central limit theorem* has given statistics its focus on averages; or because the theories for extremes are less familiar, less unified, and more complex. Moreover they are more recent and some aspects of these theories are not yet fully developed.

Let us now see how extreme values are analyzed. Extreme values are usually analyzed using one of the three major theories;

- (1) The Extreme Value Theory that deals with “annual” maxima or minima.
- (2) The Threshold Theory that deals with values above or below a specified threshold.
- (3) The Theory of Records that deals with values larger or smaller than all the previous values.

These theories deal with the actual values of the extremes. The frequencies of extremes are analyzed using the Theory of Exceedances. This theory deals exclusively with number of times a chosen threshold is exceeded. For example, in track and field, athletes may be required to meet or exceed a pre-specified performance levels in order to qualify for major competitions. Here we may be concerned with the number of the athletes that succeed. Alternatively, we may be interested in a level that ensures a certain number of qualifiers. Note that, in most applications, the number of exceedances and values of excesses over a threshold are combined to yield a more detailed analysis. For example, insurance companies analyze both the number of times a “large” claim is made and the amount by which these claims exceed a specific large threshold. Reliability engineers study both the number of extreme loads, such as large earthquakes, and their magnitudes. In what follows a brief account of each theory is given.

Extreme Value Theory

Extreme value theory generally deals with the annual (or any other period) maxima or minima. See e.g. [1] [2] [3]. Specifically, the theory is based on dividing the sample into a sub-samples and fitting a distribution to maxima or minima of the sub-samples. For example data may consist of largest earthquakes in California for each of the last 100 years.

Here like most statistical theories, first distribution of the largest or the smallest values were derived for a finite sample. Then by letting sample size tend to infinity the limiting distribution of extreme values were obtained. For this a typical maxima Y_n is reduced with a location parameter β_n and a scale parameter α_n (assumed to be positive) such that the distribution of standardized extremes $(Y_n - \beta_n) / \alpha_n$ is non-degenerate. The forms of the limiting distributions are specified by the extreme value theorem. This theorem states that there are three possible types of limiting distributions (denoted by $F_Y(y)$) for maxima. They are;

1. The Gumbel distribution (Type *I*) for which

$$F_Y(y) = \exp(-\exp(-y)) \quad \text{for } -\infty < y < \infty$$

2. The Frèchet distribution (Type *II*) for which

$$F_Y(y) = \begin{cases} 0 & \text{for } y \leq 0 \\ \exp(-y^{-k}) & \text{for } y > 0 \quad (k > 0) \end{cases}$$

3. The Weibull distribution (Type *III*) for which

$$F_Y(y) = \begin{cases} \exp(-(-y)^k) & \text{for } y \leq 0 \quad (k > 0) \\ 1 & \text{for } y > 0 \end{cases}$$

These three forms can be combined to yield the generalized extreme value distribution taking the form

$$P(\max(Y_1, Y_2, \dots, Y_N) \leq y) = \exp\{-\lambda(1 - ky / \sigma)^{1/k}\}$$

where depending on whether parameter k is positive, zero, or negative we get type *I*, type *II*, or type *III* extreme value distribution respectively.

Most classical distributions fall in the domain of attraction of one of these three types. For example distribution of maxima of samples from a normal distribution tends to Gumbel distribution. Gnedenko [4] has given the necessary and sufficient conditions for a particular distribution to fall in domain of attraction of one of the three types. For Type *I* it is

$$\lim_{t \rightarrow +\infty} n[1 - F_Y(\alpha_n t + \beta_n)] = e^{-y}$$

For Type *II* it is

$$\lim_{t \rightarrow +\infty} \frac{1 - f_Y(t \cdot y)}{1 - f_Y(t)} = y^k$$

where $t > 0$ and $k > 0$ and $f(\cdot)$ denotes the density function. Finally, for Type *III* it is

$$\lim_{t \rightarrow 0} \frac{1 - F_Y(t \cdot y + u)}{1 - F_Y(t + u)} = y^{-k}$$

where u is the endpoint of the distribution for Y ($F_Y(u) = 1$), $t > 0$, and $k < 0$.

Basic results obtained by in [4] are:

- Only distributions unbounded to the right can have a Fréchet distribution as a limit.
- Only distributions with finite right end- point ($u < \infty$) can have a Weibull as a limit.
- The Gumbel distribution can be the limit of bounded or unbounded distributions.

How do we decide which of the three limiting distributions to fit to the data? Theoretically, we can use the fact that each of the classical distributions falls in the "domain of attraction" of one of the limiting distributions above. This works if distribution of the original data is known. Unfortunately, in practice such information is not usually available and decisions should be based on the area of application or on expert opinion. For example, in the case of sports one may think of a possible ultimate record, in which case the distribution bounded above (below) is more appropriate. When information about the appropriate limiting distribution is absent, statistical goodness-of-fit may be used.

When fitting extreme value distributions to annual maxima or minima, it is possible to discard some relevant data related to the years with several large observed values and retain less informative data from the years with no real large values. The threshold theory discussed next avoids this problem.

Threshold Theory

The threshold theory allows one to make inference about the values above or below a threshold, that is, the upper or the lower tails of a distribution. It considers the excesses, the differences between the observations over the threshold and the threshold itself. See e.g. [5,6,7]. Like extreme value distributions, there are three models for tails. They are

1. Long tail Pareto.
2. Medium tail exponential.
3. Short tail distribution with an end point.

Again most classical distributions fall in domain of attraction of one these tail models. It has been shown that the natural parametric family of distributions to consider for excesses is the Generalized Pareto Distribution (*GPD*)

$$H(y; \sigma, k) = 1 - \left(1 - \frac{ky}{\sigma}\right)^{1/k}$$

Here $\sigma > 0$, $-\infty < k < \infty$ and the range of y is $0 < y < \infty$ ($k \leq 0$), $0 < y < \sigma/k$ ($k > 0$).

This is motivated by the following considerations.

- The *GPD* arises as a class of limit distributions for the excess over a threshold, as the threshold is increased toward the right-hand end of the distribution, i.e., the tail.
- If Y has the distribution $H(y, \sigma)$ and $y' > 0$, $\sigma - ky' > 0$, then the conditional distribution function of $Y - y'$ given $Y > y'$ is $H(y; \sigma - ky', k)$. This is a 'threshold stability' property; if the threshold is increased by an arbitrary amount y' , then the *GPD* form of the distribution is unchanged.
- If N is a Poisson random variable with mean λ and Y_1, Y_2, \dots, Y_N are independent excesses with distribution function $H(y; \sigma, k)$, then

$$P(\max(Y_1, Y_2, \dots, Y_N) \leq y) = \exp\{-\lambda(1 - ky/\sigma)^{1/k}\}$$

which is the generalized extreme value distribution. Thus, if N denotes the number of excesses in, say, a year and Y_1, Y_2, \dots, Y_N denote the excesses, then the annual maximum has one of the classical extreme value distributions.

- The limit $k \rightarrow 0$ of the *GPD* is the exponential distribution.

In practice the proposed method is to treat the excesses as independent random variables and to fit the *GPD* to them. The choice of threshold is, to a large extent, a matter of judgment depending on what is considered large or small.

The theory is very useful when modeling large values based on observed large values is of main concern. Clearly the modeling and prediction of large earthquakes should be based on past large earthquakes not on past medium or small earthquakes. The same is true for sports records and performances as moderate values do not carry information about the exceptional future performances.

Theory of Records

Theory of records deals with values that are strictly greater than or less than all previous values. See e.g. [8] [9] [10]. Suppose that data consists of the real numbers Y_1, \dots, Y_n with Y_n representing the most recent measurement. Usually Y_1 is counted a record, as it is the largest value at the starting point. Y_i is a record (upper record or record high) if it is bigger than all previous values or measurements. That is if

$$Y_i > \max(Y_1, \dots, Y_{i-1}) \quad \text{for } i \geq 2$$

The study of such values, their frequency, times of their occurrence, their distances from each other, etc. constitutes the theory of records. Formally the theory of records deals with four main random variables:

- (1) The number of records in a sequence of n observations.
- (2) The record times.
- (3) The waiting time between the records.
- (4) The record values.

It is interesting to note that the first three can be investigated using non-parametric methods whereas the last one requires parametric methods.

Records in general and sports records in particular are of great interest and their occurrence usually results in a great deal of media attention. Examples include the chase of the single season home run record by baseball players and breaking of the men's 100 meter record.

Conclusion

Traditionally, data values with high frequencies and averages have been the focus of statistical analysis and modeling. This has rested on the important and powerful *central limit theorem*. Extremes, rare events, and records have been neglected – treated as outliers rather than as important information. In this article, we have cited examples that point to the importance of these non-average values in the modern technological world. We think that there is a need to introduce students to these vital concepts for, not only are they important; but also the theory surrounding them can be simplified for introduction to the beginners.

REFERENCES

- [1] Gumbel, E. J. , *Statistics of Extremes*. New York: Columbia University Press, 1958.
- [2] Castillo, E., *Extreme Value Theory in Engineering*, London: Academic Press, 1987.
- [3] Coles, S. *An Introduction to Statistical Modeling of Extreme Values*, London: Springer-Verlag, 2001.
- [4] Gnedenko, B.V. “Sur la distribution limite du terme maximum d’une série alatoire” *Annals of Mathematics*, 1943, 44, 423-453.
- [5] Pickands, J., “Statistical inference using extreme order statistics,” *Annals of Statistics*, 1975, 3, 119-130.
- [6] Smith, R. L. , “Estimating tails of probability distributions” *Annals of Statistics*. 1987, 15, 1174-1207.
- [7] Anděl, J., *Mathematics of Chance*, New York: John Wiley and Sons, 2001.
- [8] Ahsanullah, M., *Record Statistics*, Commack, New York: Nova Science Publishers, 1995
- [9] Arnold, B., Balakrishnan, N., and Nagaraja, H.N., *Records*, New York: John Wiley and Sons, 1998.
- [10] Gulati, S. and Padgett, W.J. *Parametric and Nonparametric Inference from Record- Breaking Data*, New York: Springer-Verlag, 2003.

THE IMPACT OF SMART BOARD TECHNOLOGY SYSTEM USE ON STUDENT LEARNING, SATISFACTION, AND PERFORMANCE

Stuart H. Warnock, Tarleton State University (254) 968-9333, warnock@tarleton.edu
Nancy J. Boykin, Tarleton State University, (254) 968-9054, boykin@tarleton.edu
Wei Chih Tung, Texas A&M University

ABSTRACT

Responding to demands from tech-savvy students and pressures to increase student learning, faculty are incorporating instructional technology into their classrooms. Literature on instructional technology widely touts its potential for increasing student learning, satisfaction, and performance. But are these benefits universal and do they apply to all contexts and forms of instructional technology? This study focuses on the Smart Board Technology System (SBTS) and its impact on students. Responses from 111 students in a College of Agriculture and Human Sciences at a public university in the Southwest United States yield mixed results. SBTS use is shown to be positively related to both student learning and satisfaction, but not performance.

INTRODUCTION

The use of instructional technology in the classroom such as wireless internet access, presentation software, “smart boards,” real-time response systems, etc. are providing a media rich learning environment that the technologically savvy college student expects (O’hara & Seeman, 2006). According to Milliron (2001), seventy-five percent of entering traditional freshmen have substantial experience accessing and utilizing information technology. Given the primacy of information technology in students’ lives, educators have the responsibility to examine how their use of IT, or lack thereof, effects student outcomes (Fox, 1999; Green, 1999; Upcraft & Terenzini, 1998).

RELATED LITERATURE

Previous research indicates suggest that the use of information technology can result in a variety of positive outcomes such as student satisfaction (Turman & Schrodt, 2005), enhanced learning (Althaus, 1997) and significantly higher performance (Alavi, 1994; Rutz, 2003). Therefore, it is important that all these outcomes be examined when determining the appropriate use and potential impact of instructional technology in the classroom environment (Flanagin, 1999; Witmer, 1998; Lane & Shelton, 2001).

Student Satisfaction

According to Oliver & DeSarbo (1989), “student satisfaction refers to the favorability of a student’s subjective evaluation of the various outcomes and experience associated with education.” It has also been found that if the environment of the classroom correlates with the preferences of the students, then satisfaction with the educational experience will occur (Fraser, 1994).

Student Learning

Affective learning is defined as the student's emotional response to the professor, the course content and the learning environment and has been found to be related to the professor's mode of communication and instruction (Kearney, 1994). Educators and students alike believe that utilization of instructional technology facilitates learning and the ability to apply knowledge in an analytical manner (Alavi, 1994). There is an abundance of literature related to the use of specific types of technology and its relationship to enhanced learning (Arbaugh, 2000; Hall et al., 2003; Perry & Perry, 1998; Reinhardt, 1999).

Student Performance

In the literature focusing on the use of instructional technology and student performance, final course grades typically are used as the measure of performance. Alavi (1994) used two comparison groups to determine the impact of computer mediated courses on final course grades. Students in multiple sections of the same course who were exposed to computer mediated learning earned significantly higher final grades. Rutz (2003) also evaluated the utilization of various forms of instructional technology and concluded "that time on task and interest in content" were improved when instructional technology was used in the classroom, leading to higher student performance.

PURPOSE OF THE STUDY

The purpose of this study was to investigate whether a relationship exists between faculty's use of the Smart Board Technology System (SBTS) in the classroom and student outcomes. The SBTS is a multimedia instructional technology that consists of an interactive smart board screen, a data projector, and a faculty operated multimedia desk. Faculty can use their finger or a pen to reorganize nodes in a model, illustrate relationships between concepts, or mark up documents in Excel, PowerPoint, NetMeeting, etc. Using the SBTS, faculty users can incorporate and mark up still or moving images from DVD, VCR, document cameras, and computer files; content that can be saved and later exchanged.

This study explored the impact of SBTS use on the satisfaction, perceptions of learning, and performance outcomes of university students. The specific research question addressed by this study is: Does a professor's use of Smart Board technology learning environment effect student outcomes? The following three related hypotheses were tested.

H1: Is extent of Smart Board use associated with student satisfaction with the course?

H2: Is extent of Smart Board use associated with student perceptions of learning in the course?

H3: Does student performance differ based on the extent of Smart Board usage in the course?

METHODOLOGY

The sample consisted of students, graduate and undergraduate, enrolled in courses in a College of Agriculture and Human Sciences at a moderately-sized public university in the Southwest United States. Based on previous survey results concerning faculty experiences with the SBTS, fifteen faculty representing a continuum of SBTS use were selected as the primary "clusters" for this study. To be more precise, the Spring 2007 courses offered by these faculty served as the primary clusters.

The Survey

An original twenty-one item survey instrument was designed for purposes of data collection. Three multiple-item scales were used to measure the three constructs of: (1) *extent/frequency of SBTS use* – three original items, (2) *perceptions of learning* - six items modeled after Alavi (1994), and (3) *satisfaction with the course* - five items modified from Arbaugh (2000). Two items were used to measure *student performance* and the remaining items were demographic items.

The survey items consisted of a mixture of five-point Likert scales, category scales, and simple dichotomy scales. To mitigate mono-method bias, some of the multiple measurement items on each scale were reverse scaled and the ordering of the multiple measurement items was randomized throughout the survey.

Data Collection

The participation of the selected faculty members was solicited both in person by one of the authors and through a memorandum from the dean of the college. Each faculty member was asked to distribute a memo to the students in his or her class. The student memo explained the purpose of the study and invited the student respondents to visit a website where they could complete an online survey. The survey began with an operational definition of the SBTS and the student respondents were promised complete anonymity. Data were collected during the last three weeks of the Spring 2007 semester so that student respondents would be able to reliably estimate their final course performance (i.e., anticipated course grade).

Data Analysis

Respondents were profiled on all demographic variables through the use of frequency distributions. The derived multiple item measurement scales were purified via item-scale correlation and reliability analyses to provide evidence as to their construct validity.

RESULTS

It is difficult to estimate a response rate for the survey given that some students were enrolled in multiple clusters/courses, but 111 usable surveys were completed. Table 1 profiles respondent demographics.

Scale Purification

It was necessary to purify the multiple-item measurement scales. First, summated scales were created by averaging the multiple items assigned to each measurement scale. One item on the satisfaction scale was deleted because it exhibited very high correlations with more than one of these summated scales. After this deletion, all remaining measurement items showed the desired pattern of correlating with their intended scale to a greater degree than any alternative scale, all by a wide margin.

Next, scale reliability coefficients (i.e., Cronbach's α) were computed. The three-item extent/frequency of SBTS use scale (FREQUSE) and four-item course satisfaction scale (SATIS) exhibited good reliability ($\alpha = 0.73$ and $\alpha = 0.76$ respectively). One item from the six-item perceived learning (LEARN) scale was deleted in order to improve scale reliability and the resulting five-item scale shows excellent reliability ($\alpha = 0.92$).

Table 1 – Sample Demographics

<i>Characteristic</i>	<i>Levels</i>	<i>Number</i>	<i>Percentage</i>
Gender	Male	47	42.3
	Female	64	57.7
Class	Freshman	10	9.0
	Sophomore	7	6.3
	Junior	44	39.6
	Senior	45	40.5
	Graduate	5	4.5
Age	19 or under	12	10.8
	20-21	38	34.2
	22-23	36	32.4
	24-25	13	11.7
	26 or over	12	10.8
Student Type – self described	F student	0	0
	D Student	0	0
	C Student	7	7.2
	B Student	72	64.9
	A student	31	27.9
Anticipated Course Grade	F	0	0
	D	1	0.9
	C	17	15.3
	B	42	37.8
	A	51	45.9

Frequency of SBTS Use

As noted before, responses were averaged on the three items comprising this scale to create a summated scale (FREQUSE). Table 2 profiles the participating faculty concerning their extent/frequency of SBTS use. Based on responses to a former survey of the selected faculty concerning their use of the SBTS, faculty were classified a-priori into “high SBTS use” and “low SBTS use” groups. To assess the criterion validity of the FREQUSE scale, a t-test was performed and the average FREQUSE score was found to be significantly ($p = 0.01$) greater for the “high SBTS use” versus the “low SBTS use” group.

Table 2 – Rated Extent/Frequency of SBTS Use by Faculty

<i>Average Scale Value</i>	<i>Descriptor</i>	<i>Number</i>	<i>Percentage</i>
1 - 1.99	Very Infrequently	4	4.0
2 - 2.99	Infrequently	9	8.9
3 - 3.99	Neither Frequently nor Infrequently	34	33.7
4 - 4.99	Frequently	38	37.6
5	Very Frequently	16	15.8

Frequency of SBTS Use and Perceived Learning

A stepwise multiple regression model was constructed to model the variation in student perceptions of learning (LEARN) as a function of the extent/frequency of SBTS use (FREQUSE) using demographics of age, gender, and student type (i.e., STUTYPE = A student, B student, etc.) as control variables. The final model retained only one independent variable (FREQUSE), is highly significant ($F=56.2$, $p=0.00$), and explains a sizable thirty-seven percent of the variance ($r^2=0.37$) in student perceptions of learning. Collinearity diagnostics indicate no particular concerns. Table 3 summarizes the model output. Note that extent/frequency of SBTS use (FREQUSE) is highly related to student perceptions of learning (LEARN) but that none of the demographic variables have any predictive value.

Table 3 – Student Perceptions of Learning (LEARN) Regression Model

Variable	Beta Coefficient	<i>t</i> -statistic	p-value	VIF
INTERCEPT	0.953	5.187	0.000	N/A
FREQUSE	0.614	7.496	0.000	1.000
AGE	0.100	1.234	0.220	1.000
GENDER	-0.106	-1.310	0.194	1.014
STUTYPE	-0.042	-0.516	0.607	1.004

Frequency of SBTS Use and Satisfaction

A stepwise multiple regression model was used to examine the variation in student satisfaction with the course (SATIS) as a function of the extent/frequency of SBTS use (FREQUSE) and student perceptions of learning (LEARN), using age, gender, and anticipated course grade (ANTGRADE) as control variables. The final model retained two independent variables (LEARN and FREQUSE), is highly significant ($F=49.5$, $p=0.00$), and explains a substantial fifty-one percent of the variance ($r^2=0.51$) in student satisfaction with the course. Collinearity diagnostics again indicate no particular concerns. Table 4 summarizes the model output. Note that both perceptions of learning (LEARN) and the extent/frequency of SBTS use (FREQUSE) are significantly related to student satisfaction with the course (SATIS). From a practical perspective, perceptions of learning (LEARN) is a much more substantive contributor to satisfaction as its Beta coefficient is about three times that for the extent/frequency of SBTS use (FREQUSE).

Table 4 – Student Satisfaction (SATIS) Regression Model

Variable	Beta Coefficient	<i>t</i> -statistic	p-value	VIF
INTERCEPT	0.737	4.788	0.000	N/A
LEARN	0.480	6.519	0.000	1.548
FREQUSE	0.164	2.157	0.034	1.548
AGE	-0.035	-0.479	0.633	1.015
GENDER	-0.002	-0.024	0.981	1.040
ANTGRADE	-0.023	-0.299	0.765	1.151

Frequency of SBTS Use and Performance

Former studies have found a link between the use of instructional technology and student performance in the form of course grades (e.g., Alavi, 1994). This study collected two variables related to student performance. The first variable, anticipated course grade (ANTGRADE), is a weak surrogate for performance given the obvious restriction of range arising from the clear grade inflation reflected in the data of Table 1. Moreover,

there are so many intervening variables that would influence a student's course grade that using a simple course grade or anticipated course grade is a flawed approach at best. One of the most profound influences on a student's grade is the student's native intellect or academic ability. This study used a second self-report item that asked students to categorize themselves as an A-student, B-student, etc. (STUTYPE). This variable may be considered a surrogate for the student's long-term academic ability or potential.

The approach to measuring student performance took a disconfirmation/confirmation approach. A difference score (PERFDIF) between the student's anticipated course grade (ANTGRADE) and long-term academic ability or potential (STUTYPE) was computed. In instances where this difference score was negative (i.e., $PERFDIF < 0$), a *negative disconfirmation* occurred; the student under-performed their academic potential. If the student met their potential in the course (i.e., $PERFDIF = 0$), then a *confirmation* of their academic potential occurred. If the difference score was positive (i.e., $PERFDIF > 0$), then the student experienced a *positive disconfirmation*; the student over-performed their academic potential.

Given that the extant literature concerning instructional technology emphasizes its potential for *positively impacting* learning and student performance, negative disconfirmations were excluded from further analysis. The focus was left on two groups; students having a confirmation of their academic potential versus those experiencing a positive disconfirmation. If the literature is correct, then it would stand to reason that extensive/frequent instructional technology (SBTS in this study) use should lead to positive disconfirmations as students' learning is enhanced by the application of instructional technology.

A t-test was conducted using the extent/frequency of SBTS use (FREQUSE) as the response variable and PERFCNF as the factor at two levels -- confirmation versus positive disconfirmation. Equality of variance among the levels of the factor (PERFCNF) is a critical assumption that was tested finding no significant difference between the two levels ($p = 0.16$). Unfortunately, no significant difference was found in the extent/frequency of SBTS use (FREQUSE) reported by students experiencing positive disconfirmations versus those experiencing a confirmation of their academic potential ($t = 0.61$, $p = 0.54$).

CONCLUSION AND FUTURE DIRECTIONS

The results of this study confirm the findings in the literature that the use of instructional technology is positively related to student learning and satisfaction. Use of the Smart Board Technology System (SBTS) was found to be positively associated with student perceptions of learning that occurred in the course and with student satisfaction with the course. Contrary to the literature, student performance did not appear to be related to the extent/frequency of SBTS use.

This contrary result is perhaps an artifact of the weakness of this study's measure of student performance. Yet, this study did not stray far from the approach used by other researchers. The authors are considerably skeptical about this particular finding in the literature. More work is needed wherein a much more robust measure of student performance is used. Perhaps research involving a (1) narrowly defined and delimited performance task, coupled with a (2) clearly defined and robust performance rubric to measure performance would rectify this limitation of the current study and the extant literature. The authors were unable to find research that operationalized student performance in such a systematic manner. The literature could clearly be enhanced by such an approach to future research.

References available upon request to warnock@tarleton.edu.

USING WIKIS, PODCASTS, AND VIDEO FILES TO ENHANCE THE LEARNING EXPERIENCE

Steven Tello, University of Massachusetts Lowell, 978-934-4020, Steven_Tello@uml.edu
David Lewis, University of Massachusetts Lowell, 978-934-2758, David_Lewis@uml.edu

ABSTRACT

Today's college students are well versed in the use of new Internet and WWW technologies. Many of these technologies fall under the umbrella of Web 2.0. A key differentiator of Web 2.0 is an emphasis on the social creation and sharing of information. This paper examines a pilot project integrating Web 2.0 tools into undergraduate business classrooms. The authors share both successes and challenges in introducing a range of Web 2.0 tools into their classes.

Keywords: Web 2.0, classroom technology, teaching

BACKGROUND

Over the past 5 years, emerging Web 2.0 technology has become increasingly available to those with Internet and WWW access. While new software tools have supported the development of Web 2.0 tools such as wikis, blogs and podcasts, perhaps the greater shift has occurred in how people use the Internet and WWW in Web 2.0 [2] [4] [6] [7]. Batson suggests Web 2.0 represents a shift from earlier personal productivity applications which allow an end user to complete a task or job on their individual desktop to the sharing and placement of information in a social context accessible to all with Internet access [1]. Many faculty use online course management systems as a course organizer (e.g., to share lecture notes, collect homework assignments, conduct online exams), an online course administration tool geared toward increasing productivity inside, and outside, the classroom [3]. However many of our students are using Web 2.0 to share their experiences and perspectives globally using a variety of media-rich communication tools which can be accessed on their laptop, cell phone or MP3 player [6].

Today's college students are well versed in the use of new, portable technologies. A 2007 survey of 18- to 24-year old college students enrolled in four-year colleges or universities found that 97% had a cell phone, 79% had a laptop computer and 73% had a media player of some sort [4]. Some authors suggest that faculty may in fact be operating at a technology disadvantage, having adopted older desktop and wired technologies, while our younger students have come of age using multi-function mobile technologies for educational and social purposes [5]. What sets many of these newer, mobile technologies apart is their integration with Web 2.0 technology, and the younger generations embrace of the social computing paradigm supported by Web 2.0 [2] [4].

The need to better understand how our students perceive the use of Web 2.0 in our courses, as well as to gain a professional understanding of which tools should be used to teach which lessons served as the major impetus for this project.

PILOT PROJECT

In June 2007, the authors received funding under a campus educational technology grants program to pilot the use of Web 2.0 tools in several undergraduate business classes. The pilot project commenced during the Fall 2007 semester, with a small selection of Web 2.0 tools introduced into three undergraduate classes (i.e., Leadership Processes, Managing Innovation, Management Information Systems). The tools used varied based on the course and assignment. A summary of Web 2.0 tools used during the Fall 2007 semester is included in Table 1 and a lengthier discussion follows below.

TABLE 1

Web 2.0 Tools used in Pilot Semester

Class	Web 2.0 Tool	Purpose
Leadership Processes N=45	Video Podcasts 2 sessions	Exam review lectures
	Camtasia Flash Files 2 sessions	Lecture review
Management Information Systems N=35	Wimba Voice Chat 1 session	South African guest lecturer to class
	YouTube Video 6 sessions	
Managing Innovation N=29	Video Podcasts	Student review of their class presentations
	Camtasia Flash Files 5 sessions	Student review of guest lecturers
	2 sessions	Exam review lectures

The pilot semester was used to familiarize the authors with the technologies and their use in the classroom. In addition to the Web 2.0 tools noted above, each course used Blackboard Vista as a course management and communications tool, allowing each instructor to share pertinent course information (e.g., syllabus, assignments, lecture notes, assignments) with students in each face-to-face course. An end-of-semester survey was administered to students in two of the three pilot semester classes (i.e., Leadership Processes, Managing Innovation) to assess student use of Web 2.0 tools in these classes along with the students' perceptions regarding the contribution each tool made. The third course (i.e., Management Information Systems) utilized a focus session format to discuss students' perceptions of the usefulness of each tool. This data is summarized below.

THE TOOLS

Social Media – Podcasts

Podcasts refer to compressed audio files that are saved on a central website and then downloaded by end users for use on their desktop and/or portable MP3 player. A unique feature of podcasts is the ability to subscribe to the podcast, meaning end-users request that any additional podcasts from a site be automatically sent to the end user when they next launch their podcast software. Podcast software such as iTunes, IPodder, and Doppler allow end users to manage their

subscriptions, files and synchronization with portable MP3 players. An increasingly popular variant of the podcast is the video podcast, which includes both an audio track and video image. Video podcasts can be used to present a lecture, matching the audio lecture to images or slides.

The authors used podcasts in two different ways. They both created video podcasts of course lectures and exam reviews for review by students outside of class as well as integrating the review of podcasts created by others into homework and course assignments. Specific uses are described below:

Video Podcast Lectures - All three courses presented students with the option of reviewing video podcasts of selected course lectures while students in two classes also had access to exam review video podcasts. Video Podcasts were created using the multimedia authoring application Camtasia. Lecture time varied from 15 minutes to 50 minutes. Anecdotal comments by students regarding the use of course lecture video podcasts were mixed. Some students found the audio stream useful, but the reduced size lecture slides were not recognizable on their MP3 screen. Students also stated they prefer to have the lectures divided into briefer 15 minute sessions rather than longer 50 minute lectures. Students indicated this was due primarily to the limited ability to index and scroll through audio and video podcasts.

In the two classes that utilized video podcasts as an exam review tool (Managing Innovation, Leadership Processes), 87.3% of students (N=64) reported using these files at least once before each of the two courses examines. In these same two classes, 88.7 % of students indicated agreement with the statement “Review of course website materials prior to each exam contributed to my learning in this course.”

Podcast Reviews - Students in Management Information Systems were required to find and review a podcast on an Information Systems topic, and then summarize the content in a 5 slide PowerPoint presentation. Students were randomly selected to present their findings to the class. Students reported “liking” this assignment since it provided a way of obtaining additional information on a topic of interest to them individually. Most students reported listening to the topic on their laptops or desktops, rather than uploading the podcast to their MP3 players.

Podcast Perspectives – A structured assignment examining ethics in Leadership Processes provided students with links to podcast interviews with a retired congressman from two competing news organizations (in this case NPR and Fox News). Students then responded to a series of questions and were asked to summarize the differences between the two perspectives. Informal feedback from students suggested that they did not typically look for opposing viewpoints when conducting course research. Several written responses to the assignment indicated that the tone of the interviewer and subject influenced their response to questions.

Social Media - Video Files

A key aspect of Web 2.0 is the sharing of self-produced content. Sites such as YouTube host and serve self-produced video clips while sites such as Flickr, SmugMug and Webshots host and serve digital photos. These sites typically charge no fees and require no special software applications. Amateur (and professional) photographers and videographers can upload, tag

(label with keywords) and share their work with anyone who has access to the Internet and WWW. Streaming video clips linked from the social networking site YouTube were integrated into the introductory Information Systems course lectures. These files were used during lectures to highlight specific course points. The video content was produced by amateur videographers not related to this course or institution.

While YouTube and similar sites provide search and review options, the first challenge confronted in using others' video clips in class is the amount of time required to find and review clips appropriate to the course content. One author spent approximately 15 hours searching for appropriate video clips, then used six whose total elapsed time was 14.5 minutes. Since tags, or keyword identifiers, are added by the author without any enforced taxonomy, one must first learn the tagging language and then hope the author tagged the clip appropriately and sufficiently. Additionally, some sites return search request data in order of ranked popularity, meaning the more frequently a video clip is viewed, the more likely it is to appear in the first several pages of search returns. The ability for an end user to find the content he or she desires is somewhat dependent on the content author's tagging experience and the viewing trends of thousands of other Web 2.0 users.

When appropriate clips were found, the author found that these were most useful when short in length (less than 3 minutes) and humorous in nature, adding levity to a some times dry academic topic. The major downside was the upfront time to identify and review the video clips, as there was no adequate search mechanism to identify the content or the quality of the video clips.

Synchronous Conferencing Tools

Horizon Wimba is a collaborative educational technology. It allows for synchronous audio lectures and conversations, multimedia presentation and the sharing of the same desktop between multiple locations. One of the authors (Management Information Systems) used this tool to invite a guest lecturer and a question and answer presentation from a colleague in South Africa. Students reported overwhelmingly liking this classroom tool primarily because they could ask questions of an expert at a distance location. While the guest lecture added to the student experience, a significant amount of time was required to prepare the technology for the presentation. When using Wimba for a remote lecture transmission, it took 6 hours of preparation time for each participant for a 1 hour 45 minute interaction. The issues encountered primarily involved firewall settings, bandwidth restrictions and user setup. These issues could be related to institution specific network policies and firewalls.

Wikis

A wiki is a website or similar online resource which allows users to add and edit content collectively. As users add or edit pages, a log of changes to the wiki is recorded. Depending on the administration of the wiki, changes may be restricted to certain users or available to any potential contributor. Wikis are quickly gaining favor among educators who believe the ease with which this Web 2.0 tool supports student and educator collaboration is important in supporting authentic learning [2].

The College of Management at the authors' university has created their own wiki for use by faculty for course material and other college wide projects (<http://comwiki.uml.edu>). One author is using the ComWiki to share videos created for his Managing Innovation class. Thus, it is a repository of information where students in future classes can view videos from previous classes, post their own videos, and comment on the videos of their classmates.

The other author is integrating the ComWiki into his Spring 2008 online Management Information Systems course. Part of the class includes a discussion of a short novel, *The Goal*, and it is anticipated that the wiki will provide an online space where students can add their own comments and key questions related to this text. The wiki tool used for this course, MediaWiki, provides WYSIWYG editing, meaning there is no need for students to learn any markup language in order to edit the site. Progress regarding the use of this wiki by the authors and their students will be discussed at the conference.

Audio Video Screen Capture

While not strictly a Web 2.0 tool, the authors used the multimedia authoring tool Camtasia to create both video podcasts and online multimedia instructional videos. We discussed the video podcasts created in the pilot above, but several other uses of Camtasia deserve mention. These include:

Tutorials. Brief audio video tutorials were developed to provide students with hints on how to use different software products required in the Management Information Systems course. Specifically, these tutorials explain navigation of the course Blackboard companion site, and how to perform specific functions in Microsoft Access and Microsoft Excel. Once developed, Camstasia allows these files to be saved in multiple online formats (e.g., Flash movie, Streaming QuickTime movie, Microsoft AVI movie).

Enhanced Podcast. An enhanced podcast involves the synchronization of an audio file to a series of image slides. In regard to the three courses discussed above, this involved synchronizing the live recording of a class lecture to PowerPoint slides from the class. The enhanced podcast was created after the class had ended, was saved as both a Flash movie and video podcast, and was then uploaded to the respected class website. This approach was especially helpful for students who missed a specific class.

Using Camtasia to capture video clips from the net. Camtasia can be used to both capture and translate video clips downloaded from other websites. This is particularly helpful when attempting to download and play a video during a class lecture. When the material is downloaded and converted to the local computer, the instructor does not have to depend on the website and network access being available.

Short snippets before class. Camtasia was used to capture and record short video clips, snippets, which discussed various aspects of the material to be discussed in that day's class. After showing these snippets to students at the start of class, students were then quizzed on the content of the snippet later in the class. The main purpose is to engage students in the subject matter prior to class.

CONCLUSIONS

Web 2.0 technologies offer faculty new tools for working with today's technology literate student. Our pilot project has identified several practical applications of Web 2.0 technology in the classroom. Our students have quickly picked up on the nuances of this technology and, in general, expressed satisfaction with the use of these tools. The potential for expanding the use of Web 2.0 in our classes is significant, tempered only by the increased time required in materials and assignment development. As faculty expand the use of Web 2.0 tools in their courses it will be important to weigh the contributions of the technology against the time required of both instructor and student.

In order to fully realize the potential of these new technologies, faculty must consider the context in which today's college students are using Web 2.0. While earlier educational technologies have been used rather successfully to administer and organize course content and activities, the entry-level, common access platform of Web 2.0 makes it relatively easy for students and faculty to produce and share information in a variety of formats. The apparent strengths of Web 2.0 (e.g., ease of use, social creation of content, multiple forms of expression) require that faculty develop course activities that emphasize student development and sharing of course artifacts in a manner that attests to learning that occurred throughout the course. Even with these challenges, we believe that use of Web 2.0 tools is beneficial to the overall learning experience of our online and on-campus business students.

REFERENCES

- [1] Batson, T. "A new social context for information." *Campus Technology Magazine*, Jan. 16, 2008. Accessed online 1/17/2008 at <http://campustechnology.com/articles/57281/>.
- [2] Beldarrain, Y. "Distance Education Trends: Integrating new technologies to foster student interaction and collaboration." *Distance Education*, August 2006, 27(2), 139-153.
- [3] Borreson Caruso, J. "Measuring student experiences with course management systems." Educause Center for Applied Research, Research Bulletin, V 2006, 19.
- [4] Craig, E. "Changing paradigms: managed learning environments and Web 2.0," *Campus-wide Information Systems*, 2007, 24(3), 152-161.
- [5] Eduventures. "Snapshot: personal electronic devices owned by students," *Campus Technology Newsletter*, 1/8/2008, accessed online 1/11/2008 at <http://www.campustechnology.com/article.aspx?aid=57155>,
- [6] Gooding, J. "Web 2.0: A vehicle for transforming education," *International Journal of Information and Communication Technology Education*. Apr-Jun 2007, 4(2), 44-54.
- [7] McGee, P. & Diaz, V. "Wikis and podcasts and blogs! Oh, my! What is a faculty member supposed to do?" *Educause Review*, Sept./Oct. 2007, 42(5), 28-40.

USING ERP SOFTWARE TO TEACH DBMS CONCEPTS: A HANDS-ON APPROACH

Deb Sledgianowski, Department of Information Technology and Quantitative Methods,
Hofstra University, Hempstead, NY, (516) 463-4759, Deb.Sledgianowski@hofstra.edu

Michael Devlin, Office of Student Affairs, Hofstra University, Hempstead, NY,
(516) 463-4318, Michael.P.Devlin@hofstra.edu

ABSTRACT

This paper is a work-in-progress discussing our ongoing development and use of course materials to demonstrate database management systems concepts to students by using enterprise resource planning system software to augment their learning. The purpose of our paper is to share our experience with other universities interested in using ERP software as a tool to supplement students' learning of business and information technology concepts. This paper describes the course material and tutorials we developed to facilitate students' understanding using a hands-on approach. We used Oracle as our database management system and SAP R/3 as our ERP system. Experiential learning theory provides the guiding framework for our approach.

Keywords: curricula, database management, enterprise systems, teaching

INTRODUCTION

Business school accreditation standards [1] call for participating schools to continuously improve instructional programs. Information systems curricula models suggest that students have an understanding of information technology (IT) from a business perspective at the enterprise level and that students have both an applied and conceptual learning experience [9] [5] [7]. To facilitate our students learning, we are developing hands-on learning tutorials using Enterprise Resource Planning (ERP) software to demonstrate database management systems (DBMS) administration concepts as they are applied to a popular enterprise system.

ERP software has been widely adopted by large companies and is increasingly being adopted by small to medium sized companies. It is very likely that business students majoring in information technology will be involved with ERP software in their careers as IT professionals or managers. As a business school, it is our goal to provide students with the skills and knowledge that they will need to manage in today's global business environment. The SAP University Alliance (UA) supports this goal by providing the use of their mySAP Business Suite software for a nominal fee to member schools. This software application can be used to demonstrate how IT is critical to the efficient performance of business processes that lead to increased productivity and corporate growth. A goal for our students is for them to gain hands-on experience of tasks similar to what they may experience in the real world to supplement IT concepts discussed in class.

This paper first provides a brief description of active learning theory as it relates to our hands-on practice and then describes the curriculum material we are developing and how we are using it with our students.

IMPLEMENTING ACTIVE LEARNING

We chose to use an active learning method to facilitate learning in an undergraduate database management systems course. A motivation for this course development was the model curricula for undergraduate information systems IS education [2] [3] recommendations to provide a mix of experiential learning and theory in courses. Additionally, business school accreditation standards [1] encourage participating schools to continuously improve instructional programs and to use interdisciplinary approaches. This course is offered by the information technology department within a business school, so guiding the instructional delivery decisions were the goals to integrate information technology concepts with business concepts as well as to provide students with the opportunity to apply hands-on technical concepts.

Active learning and experiential learning are terms used to describe educational techniques where students learn by doing. These learning concepts, in part, are founded by the writings of John Dewey and Kurt Lewin. Influenced by Dewey and Lewin is the work of Kolb and Fry [5], who proposed the experiential learning circle, based on a four stage model:

- (1) experience,
- (2) observation and reflection,
- (3) formation of abstract concepts,
- (4) testing in new situations.

This model of experiential learning is the basis of “many discussions of the theory and practice of adult education, informal education and lifelong learning” [7]. Research [5] [6] using active and experiential learning techniques suggest that students may benefit by having hands-on experiences as part of their education process.

Our DBMS course offerings at Hofstra University cover the development of a database application from conceptual modeling through logical design to implementation using desktop and enterprise DBMSs such as Microsoft Access and Oracle. Students learn Structured Query Language (SQL) to insert, update, delete, and query data in both of these relational DBMS.

A critical task of IT professionals responsible for database administration in organizations is to provide users of database applications with access to the data and to ensure the integrity and security of that data. This is commonly done by creating user accounts and granting access rights and privileges to those accounts. To broaden our student’ experience, we decided to show them how these concepts are applied to an ERP system, similar to what they may experience in the real world during the span of their IT careers. We are in the process of developing tutorials in SAP to demonstrate using an enterprise software application to provide database access security and to demonstrate creating and modifying physical database structures such as tables.

The first tutorial will demonstrate implementing data security by creating user accounts and granting them transaction authorization. This tutorial, relative to Kolb and Fry’s experiential learning circle model, is discussed below:

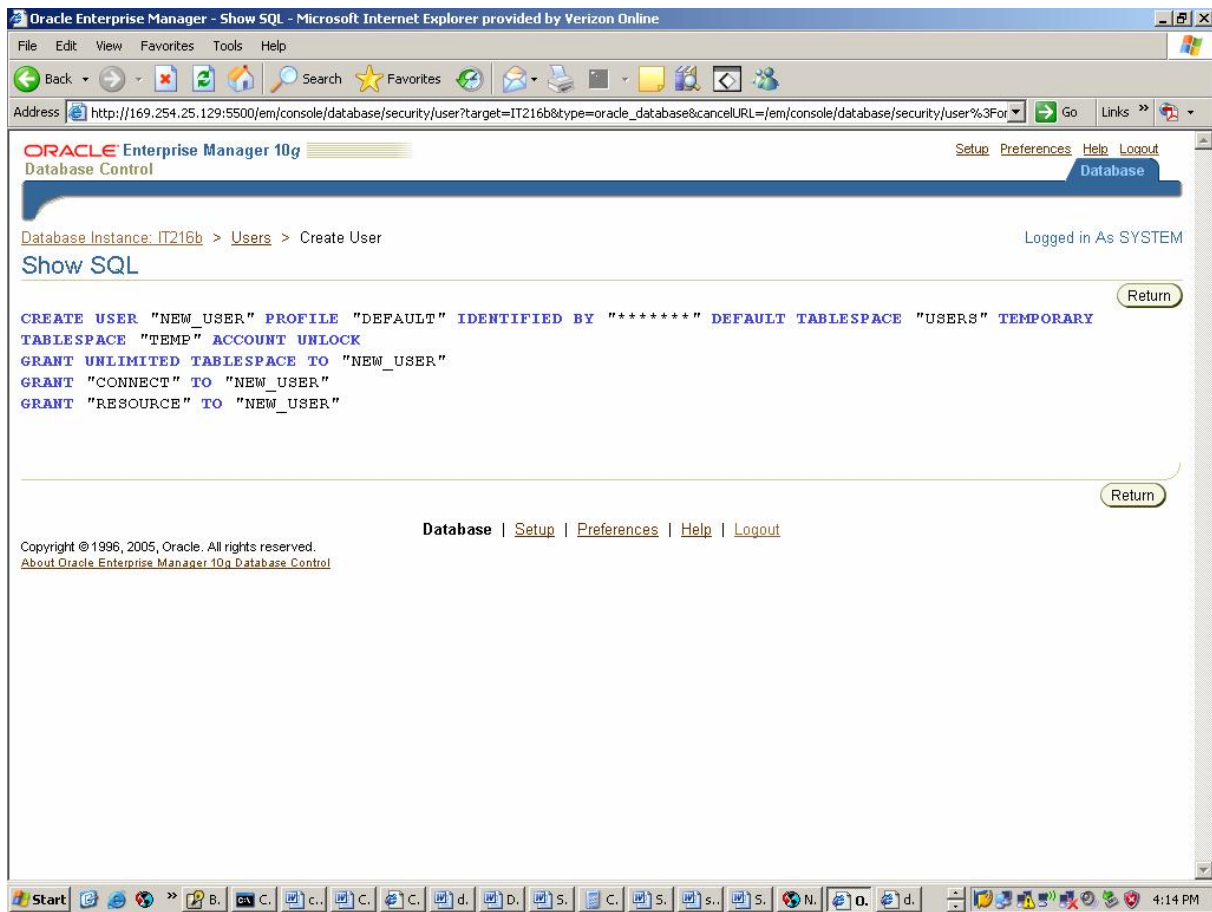
In the data security tutorial, students first learn the structure of SQL syntax and the specific commands to create users and roles and grant permissions to enable them to access database objects. We consider this to be a manifestation of the “experience” stage of Kolb and Fry’s learning cycle.

Next the students practice the SQL syntax that they just learned by executing the commands in an actual DBMS (see Figure 1 through Figure 2) – they learn to understand the effects of the SQL commands in a particular instance so that if the same actions were taken in similar circumstances it would be possible for them to anticipate what would follow from their action. We consider this to be a manifestation of the “observation and reflection” stage.

Figure 1: Creating a database role in SQL:

```
CREATE ROLE "ROLE_EXAMPLE"
```

Figure 2: SQL syntax for creating a user account with roles in Oracle:



Then the students learn how to create user accounts and grant transaction authorization by using the tools provided by the ERP system (see Figure 3 through Figure 5), the “formation of abstract concepts” stage. It is anticipated that the students come to the abstract realization that the method that they used to create the user accounts and grant access in the ERP system, in

effect caused the same SQL commands they had just learned to be executed “behind the scenes” by the database engine.

The students can then apply the concepts they have learned to new situations, the “testing in new situations” stage. In our case, we have the students learn additional SQL syntax for other DBMS functionality, such as creating and modifying physical database structures (tables and views). The students then learn how to do this with the ERP system, realizing that there are also vendor-provided tools for these tasks and that ultimately these tools execute the same SQL commands that they just learned. We anticipate this realization would also occur when the students use other ERP or enterprise systems which are built on a relational DBMS back-end.

Figure 3. Menu path to user maintenance:

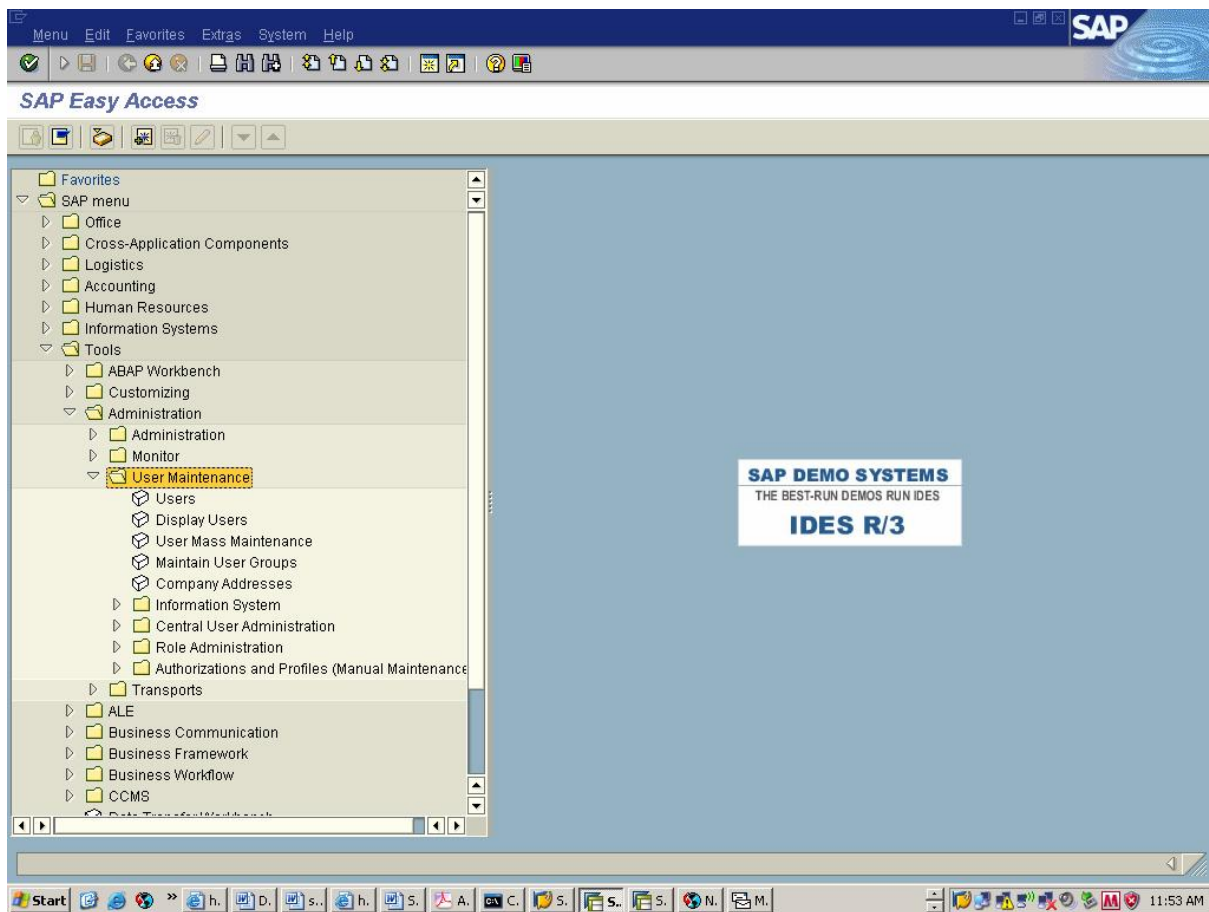


Figure 4: View, change, add user profiles:

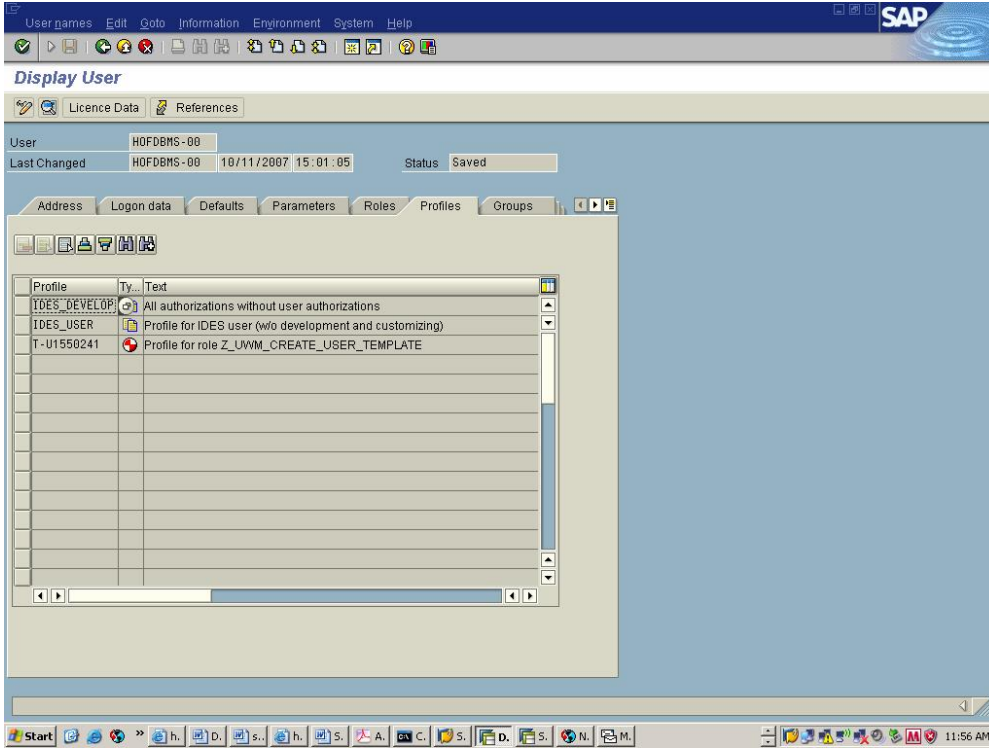
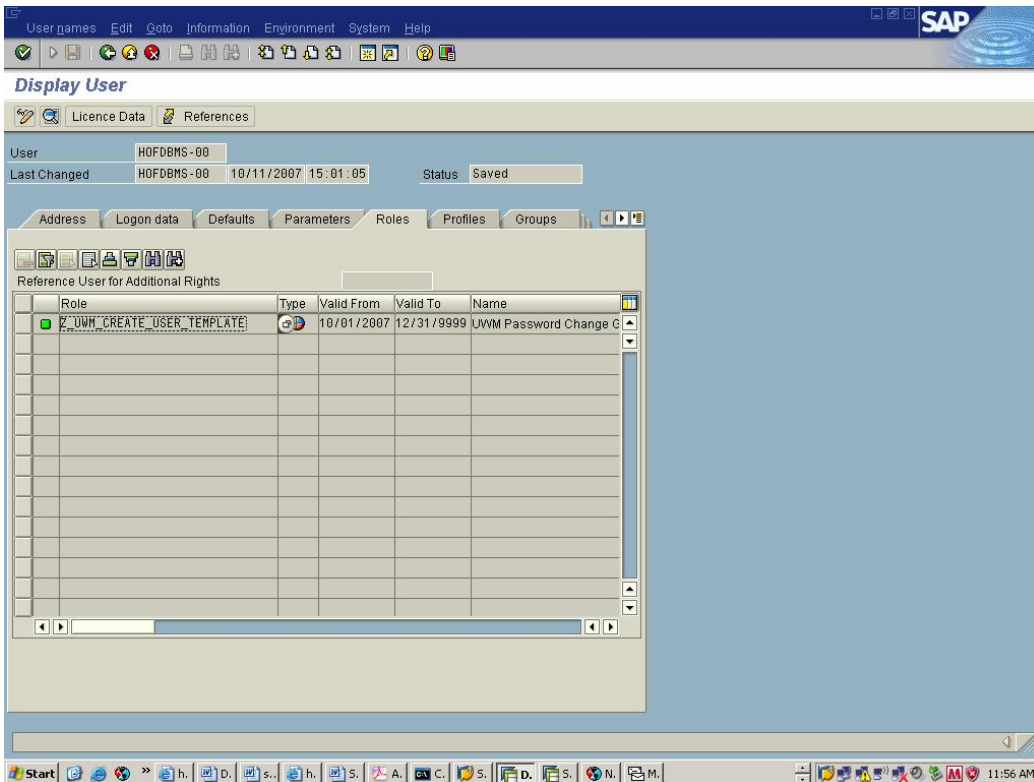


Figure 5: View, change, add roles



CONCLUDING REMARKS

At the end of the semester, students will be surveyed to assess how the course materials facilitated learning. Preliminary feedback from our students has been positive, with most believing that this experience will better prepare them in applying the concepts they learned to real world systems.

We are in the early stage of our curriculum development. Our experience with students using tutorials has been positive, encouraging us to develop additional tutorials and learning aids. Future plans include developing learning objects of DBMS concepts and publishing them in podcasts available online and distributed via our Blackboard course management system. Podcast materials can be viewed on-line or downloaded to an MP3 player. This frees the student from time constraints (they can view the materials when it's convenient to them) and from geography constraints (they can view them anywhere from a mobile device). This offloads classroom components that do not require significant student-instructor interaction, reducing the pressure on classroom time. Systematic research is needed to further examine the effects of using this new curriculum material on educational outcomes.

REFERENCES

- [1] AACSB. Eligibility procedures and standards for business accreditation. Retrieved September 20, 2007 from <http://www.aacsb.edu/accreditation/standards.asp>
- [2] Cohen, E. (ed.) Curriculum Model 2000 of the Information Resource Management Association and the Data Administration Managers Association, 2000. Retrieved October 1, 2007 from http://www.irma-international.org/downloads/pdf/irma_dama.pdf
- [3] Gorgone, J., Davis, G., Valacich, J., Topi, H., Feinstein, D., & Longenecker, H. IS 2002 - Model curriculum and guidelines for undergraduate degree programs in information systems. *Association for Information Systems*, 2002. Retrieved September 20, 2007 from <http://www.acm.org/education/is2002.pdf>
- [4] Kalkani, E., Boussiakou, I., & Boussiakou, L. (2005). The paper beam: hands-on design for team work experience of freshman in engineering. *European Journal of Engineering Education*, 30(3), pp 393-402.
- [5] Kolb, D. A. & Fry, R. (1975). Toward an applied theory of experiential learning; in C. Cooper (ed.) *Theories of Group Process*, London: John Wiley.
- [6] Marshall, K. & Pearson, M. (2007). A spreadsheet pricing model approach to teaching practical pricing concepts in the marketing classroom. *Marketing Education Review*, 17(1), pp 87-93.
- [7] Smith, M. K. (2001). David A. Kolb on experiential learning, *The encyclopedia of informal education*. Retrieved October 9, 2007 from <http://www.infed.org/biblio/b-explrn.htm>

INTERNATIONALIZING BUSINESS EDUCATION: EVIDENCES FROM LATIN AMERICA

Mohammad Elahee, Quinnipiac University, Hamden, CT 06518

Tel: (203) 582-3453; E-mail: mohammad.elahee@quinnipiac.edu

&

Mario Norbis, Quinnipiac University, Hamden, CT 06518

Tel: (203) 582-8309; E-mail: mario.norbis@quinnipiac.edu¹

ABSTRACT

Based on survey data from Latin America, this study examines the extent of internationalization of curricula in Latin American universities. The findings show that most business schools in Latin America have revamped business education so as to create a global awareness among their students. The study also identifies the areas that need further attention and resources for further internationalization of business education in Latin America.

Key Word: Internationalization, Business Curriculum, Latin America

INTRODUCTION

Internationalization of business curriculum has lately been receiving a lot of attention as evidenced increased number of papers published in the topic. Although a large number of studies focusing on various aspects of internationalization of business curricula in the context of the US have been published, we do not find much work on the same topic in the context of non-US countries. It would therefore be interesting to see what other countries are doing in this regard. Although a few studies can be found in the context of Europe and Australia, very few studies have been conducted about the state of business education in Latin America, an important business partner of the US. Even the Academy of International Business (AIB) sponsored “Sixth global Survey of Business Schools” [1], arguably the most comprehensive survey on internationalizing business education around the world, hardly mentions anything about Latin America². This study aims to partially fill this void by focusing on internationalization of business education in Latin America. For the purpose of this paper, Latin America includes all countries south of the US, except those in the Caribbean region .

Internationalization of Business Education and Latin America

To gain an understanding about the extent of internationalization of business education in Latin America, this paper examines the perception of academic administrators about the importance of different aspects of international business education and the extent of internationalization of

¹ Both authors have made equal contribution to writing this paper. Their names appear here in alphabetical order.

² This Global Survey reported that out of 151 surveys received, 102 responses came from the US, 29 responses came from Europe only 20 responses came from the rest of the world including Latin America. The study did not specify the exact number of Latin American responses.

curricula in different functional areas of business. In doing so, this paper specifically explores the following issues:

- 1) Is there any push toward internationalizing business curricula at Latin American Universities?
- 2) What are the areas that are being impacted by such push for internationalization?
- 3) What recommendations are Latin American educators making to strengthen international business education?
- 4) The extent of internationalization of business education in terms of research, faculty, overseas education (e.g., study abroad, faculty exchange etc.), and international students.

LITERATURE REVIEW

Since the seminal work of Nehrt [3] on international business education, a large number of studies have been published in this area. Based on our analysis of existing literature, we can identify four different strands in the literature on internationalization of business education. The *first strand* focuses on perceptions of educators and employers about the importance of internationalizing business education [4,5,6,7, 8]. Hofstede et al. [9], while examining the goals of business leaders, provided a long list of international business failures that resulted from a lack of education in the multi-cultural aspects of business.

The *second strand* consists of normative studies that provide “how to” directions concerning internationalizing the business education [e.g., 10, 11]. Ellingboe [12] identified five components that are integral to the internationalization processes of curriculum at universities, they are 1) college leadership, 2) faculty members’ international involvement, 3) availability of accessible study abroad programs, 4) presence of international students on campus, and 5) international co-curricular units.

The *third strand* consists of papers that studied the process of internationalization of business education in non-US universities such as those in Europe [13]; Australia [14], etc.

The *fourth strand* focuses on internationalization of business curricula within a particular discipline such as accounting [15], finance [16], marketing [17] etc and highlights the discipline specific challenges.

Previous studies [e.g., 10, 15] show that internationalization of business education is a multi-dimensional concept and these various dimensions are interrelated. Previous studies also suggest that many schools undertake the painful process of internationalizing the business education in the belief that such internationalization would result in better job placements for students, better research opportunities for the faculty, and higher reputation of the school [12, 19]. Building on such findings, this paper proposes the following:

Proposition 1: *There is a direct relationship between showing concern for placement of students and facilitating better research for faculty.*

Proposition 2: *There is a direct relationship between:*

2a. *showing concern for placement of students and the future reputation of the school.*

2b. *showing concern for faculty research and the future reputation of the school.*

METHODOLOGY

Scale Development

Data were collected by means of a self-administered survey that was sent to all the member universities of CLADEA. Following the recommendations of Churchill ([18] for developing a scale, we first identified the domain (concerns for student placement, faculty research and future reputations) and then developed a comprehensive list of initial items through reviewing the literature. Given our previous teaching and work experience in the region, we could delete many of these initial items that we felt were not appropriate in the context of Latin America. To ensure relevance, face validity, and readability with academic administrators, expert opinion about the survey was also sought two from professors who are currently teaching in Latin American universities. The final version of the scale included ten items. A few factual and open ended questions were added to the questionnaire at the end.

The instrument was translated into Spanish by the authors, one of whom is a native speaker of English and the other is also fluent in Spanish. It was verified and back translated into English by another bi-lingual faculty who teaches Spanish at a University in North-East USA.

Data Collection

For the purpose of this study, Latin America includes all countries south of the US, except the countries in the Caribbean region. We contacted the Dean/Chair of the business school of the identified universities. It may be noted here that many of the universities identified for this study do not have a business school, but simply have a department focusing on business education. We therefore included the Dean/Chair/Head of business school, faculty of commerce, department of management/administration etc. in our sample frame.

The surveys were sent to the Deans/Department chairs of all 69 business schools that are members of Consejo latinoamericano de escuelas de administracion (CLADEA), the Association of Latin America's leading management schools in self-addressed, stamped envelopes. A reminder was sent via email three weeks after the survey was mailed. We received a total of 14 responses. Later, we contacted all those who were sent the survey earlier by telephone as well as email and requested them to participate in the survey. This yielded another 7 responses, thus

raising the response rate of 31%. The overall response rate was considered satisfactory given that similar studies earlier yielded much lower response rates [e.g., 2]. Out of the 21 responses, 17 came from private universities while only four came from public universities. All but one of these responses came from universities where Spanish is the language of instructions (along with English in some cases). There was only one response that came from Brazil, stating that the language of instruction in that university was in Portuguese.

DATA ANALYSIS AND DISCUSSION

We conducted factor analysis and also computed descriptive statistics and correlations. For non-response analysis, we used *t* test. Our factor analysis yielded three factors, which we labeled as: a) concern for student placement, b) concern for faculty research, and c) concern for future reputation of the school. To assess the reliability of the data, Cronbach alphas for these factors were computed. The Cronbach alpha for these three factors were .75, .74 and .53 respectively. The level of alpha for the first two factors meet the criteria set by Nunnally [20] for descriptive studies.

The seven late responses were used as proxy for non-responses. We did not find any significant differences between the original 14 responses and the seven late responses. Our correlation results indicated a high positive correlation among the three variables. We found that the concern for student's placement was significantly correlated with concern for faculty research (87%) as well as with concern for future reputation of the school (74%). We also found a strong correlation (61%) between concern for faculty research and concern for future reputation of the school.

About 67% of the respondents stated that their universities have a department or program focusing on international business (IB). The number of faculty teaching such courses ranged from a low 1 to high 28.

43% of the respondents reported having a concentration of international business at the undergraduate level while 50% of the universities reported offering at least one course on IB. At the Master level, the number was even higher, with 64% of the institutions offering a concentration on IB while 71% of the respondents mentioned offering at least 1 course in the IB department

86% of the respondents indicated the presence of foreign students on their campus. 73% of the respondents stated that they have at least one to three faculty members with a Ph.D. in the area of international management, international marketing and international finance. 57% of our respondents indicated that they have an active study-abroad program. The range of students taking part in such study-abroad program varied from 1% to 30%, with 10% being the most typical number. The number of universities having exchange programs with foreign universities, according to our survey, was even higher. 86% of the respondents stated that their institutions have ongoing student exchange programs, while 76% reported having faculty exchange programs with foreign universities. The reason behind such a large number of faculty exchange

programs with universities in Europe is probably due to the fact that Latin American universities are reaching more to Portuguese and Spanish universities because of linguistic affinity. 76% of the respondents reported that they encourage faculty to participate in research and consulting that focuses on topics beyond their national borders, and 78% of our respondents indicated that the work of their faculty have appeared in foreign academic journals.

90% of the respondents indicated that their faculty use cases in their teaching. We also found that Latin American universities are embracing American style of business education, a fact corroborated by the number of schools that are trying to get AACSB accreditation. 81% of our respondents (excluding those that already have AACSB accreditation) stated that they are actively seeking AACSB accreditation within next five years.

CONCLUSION

While limited in scope, this study does provide some interesting insights about the current state of internationalization of business education in Latin America. It is hoped that this study would generate further interest in this topic that would eventually provide us with a greater understanding about the state of business education in Latin America and provide guidance as to how Latin American business programs can strengthen their internationalization efforts.

REFERENCES

- [1] Kwok, Chuck C.Y. and Jeffrey S. Arpan. "Internationalizing the business school: A global survey in 2000." *Journal of international business studies*. 2002, 33 (3): 571-581.
- [2] Kwok, Chuck C.Y. and Jeffrey S. Arpan.. "A global survey in 2000." *Journal of international business studies*, 1994, 25 (3):19-31.
- [3] Nehrt, L. *Business and international education*. Washington DC, American Council on Education, May, 1977.
- [4] Patrick, T. "Attitudes of corporations towards international business education." *Journal of international business studies*. 1978, 18(1): Spring, 83-90.
- [5] Philip, N. "Which international business (IB) courses do IB faculty consider the most important?" *Journal of teaching in international business*. 1995, 6(3): 87-100.
- [6] _____. "Academic-corporate sector cooperation in international business education: A faculty perspective." *Riverina Papers in Commerce*, No. 21, October, 1992, Charles Stuart University, Wagga Wagga, Australia.
- [7] Ricks, D.A. and M. Czinkota. "International business: An examination of the corporate viewpoint." *Journal of international business studies*. 1979, 10(3): Fall, 97-100.

- [8] Beamish, Paul W. and Calof, J. "International business education: A corporate view." *Journal of international business studies*. 1989, 20(3): Fall, 553-564.
- [9] Hofstede, Geert. *Culture's consequences: International differences in work related values*. Beverly Hills, 1980. CA: Sage.
- [10] Fatemi, Khosrow. "A Multi-dimensional strategy for the internationalization of business programs." *Journal of transnational management*. 1995, 19-34.
- [11] Kedia, B. and B. Cornwell. "Mission based strategies for internationalizing US business Schools." *Journal of teaching in international business*. 1994, 5 (1): 11-29.
- [12] Ellingboe, B. "Divisional strategies to internationalize a campus portrait" in J. A. Mestenhauser and B. J. Ellingboe,(Eds.), *Reforming the higher education curriculum: Internationalizing the campus*. 1998. Phoenix: Oryx Press.
- [13] Galbraith, Kate. "European business schools expand their reach." *The chronicle of higher Education*. 2002, 48(37): May 24 2002, A40-A43.
- [14] Frazer, Murray C. "Internationalizing management education: Experience in Australia." *Journal of Business Administration*. 1993, 21 (1& 2), 327-330.
- [15] Radebaugh, Lee. "Internationalization of the accounting curriculum." *Journal of business administration*, 1993, 21 (1& 2), 167-190.
- [16] Kwok, Chuck C.Y. "An empirical study of finance curriculum internationalization." *Finance practice and education*. 1994, 4(2): 35-47.
- [17] Lumstrom W., S. White, and C. Shuster. "Internationalizing the marketing curriculum: The professional marketer's perspective." *Journal of marketing education*, 1996. 18 (20): 5-16.
- [18] Churchill G.A. Jr. "A paradigm for developing better measures of marketing constructs." *Journal of marketing research*, 1979, 16(1): 64-73.
- [19] Han, Pyunge. "The international dimensions of functional business courses." *American business review*. 2002, 22(1): 56-61.
- [20] Nunnally, Jum C. *Psychometric theory*. 1st ed. 1967. New York: McGraw Hill.

FORESTALLING UNCONSCIONABLE PREDATORY FINANCIAL ADVISING PRACTICES – FINANCIAL GERONTOLOGY TEACHING NOTES

Nicole Grandmont-Gariboldi
St. Thomas University
16401 N.W. 37th Ave., Miami Gardens, Fl. 33054-6459
Tel.: 305-628-6598 Email: ngariboldi@stu.edu

ABSTRACT

This paper discusses the vulnerability of elderly persons to potential mistreatment by financial advisors. More particularly, in the context of a proposed SOWT framework, it provides the elderly with simple prevention measures and self-defense strategies against abusive financial advising practices. The framework can also be used for teaching in the area of financial gerontology. In addition, a number of scenarios are presented for discussion in the classroom.

Keywords: Financial, gerontology, abuse, strategy, prevention

INTRODUCTION

“While the phrase ‘elder abuse’ is well-known, the phrase “financial elder abuse” is not so familiar” [1]. This paper offers a framework for assessing the elderly persons’ exposure to potential abuse by financial advisors/managers and for improving their awareness and readiness to recognize and resist such threats. Self-defense strategies are proposed within the context of the proposed SOWT model. A number of scenarios are also presented for discussion in the classroom.

THE SOWT MODEL

The SOWT model presented in this paper is an adaptation of a framework developed earlier by Grandmont-Gariboldi [3]; it provides an improvement over the narrative form of the SWOT analysis (Strengths, Weaknesses, Opportunities, and Threats) often found in the current literature. *Annex 1* presents the SOWT wheel model, which involves five wheels representing respectively the strengths and opportunities on the left side [SO], the weaknesses and threats [WT] on the right side, and in the middle, the strategic actions that can be used to maximize the SO and minimize the WT. The three pillars at the basis are essential for a successful strategy. The critical part at the center of the SO wheels is the individual’s area(s) of leverage. In normal circumstances, focusing on the strengths and opportunities should allow to minimize weaknesses and threats. This is why the small circle is adjacent to the SO circles. The likely shrinking effect on the threats and weaknesses is reflected in the decrease of the size of the WT circles. The initial dimension of each circle can be determined based on quantitative and qualitative analyses of an individual’s exposure to possible mistreatment. For instance an initial likert- scale questionnaire could present a workable tool. Adjustments in the relative size of each circle can be estimated for instance using percentage changes in a person’s readiness and

willingness to resist unethical pressures. The idea is to avoid a defensive approach, which focuses on threats and weaknesses and bears the risks of developing vicious circles of actions-reactions involving the elderly person and the financial expert. However, given that we do not live in a perfect world, some weaknesses and threats may never be eliminated. Thus, they have to be included as inherent part of the process as a whole. Also, the model allows shifting the small circle to the right in particular cases, such as crisis situations.

Here in this paper we will focus on the three pillars, on the individual's area(s) of leverage and on the SO side of the model. Following is a short discussion of the basic elements expressed in terms of advice for the elderly, as presented in *Annex 1*. *Trust*: At certain points in our life we need to trust someone. Isolation is a major threat especially to an elderly person. *Family/Friends Support*: Think about your golden years as early as you can. We all need to be realistic and ready to face the inevitable ageing process and its financial consequences. Try to nurture good relationship with family or friends. Even a truly committed financial professional could be a trustworthy friend. *Financial literacy*: If you have no desire, patience, or talent to get educated on financial matters, make sure that the people who will manage your finances have the adequate credentials. If a family member or a friend is in charge of your finances, make sure that they have the minimum intellect to evaluate the qualifications of your potential financial manager. *Look for Help*: Do not try to do it all. It is possible that you are able to manage everything on your own at this point in time, but you do not know what the future holds. So you need to have a back-up plan. Also another point of view could help. *Awareness*: Do not always think of others in terms of yourself. You may have a candid and optimistic view of the people you entrust your affairs, but you need to be aware that financial professionals may not all be worthy of your trust. Beware of potential conflicts of interest or even dishonesty especially if the individual or the firm is not well established. Moreover, even in well established financial institutions, there are possible internal conflicts of interest. *Readiness*: Although you or your representative may have an adequate level of financial literacy and a certain level of awareness, the willingness to face reality and react to unethical behavior becomes crucial. You need to provide for the eventuality that your physical/psychological condition may not allow you face abusive practices. This is why it is so important to plan early and make sure that your representative will have the strength of character to take the appropriate actions. *Be receptive*: Remember that principles and practices that were adequate in the past may not be optimal today. In addition to a basic financial literacy, you or your representative needs to have an opened mind to learning. *Trust yourself*: Most of all trust your own capacity to take control to of your financial affairs, the extent that this is possible. This does not mean that you need to manage everything on your own. However, make sure that your representative is acting according to your wishes and your directives. Do not let anyone intimidate you, especially if the person who administers your affairs is also taking care of other aspects of your life, such as the choice of an assisted-living residence.

The SOWT framework proposed in this paper can be easily adapted to particular situations. It allows interactive inputs from students in the traditional or virtual classroom. In addition, the possible combination of this working tool with other models,

such as the REAL model [4], offers opportunities for creative thinking in strategic positioning.

SELECTED SCENARIOS

Following are examples of possible situations in which the elderly are ideal prey of abusive care takers and financial advisors who often use obscure and almost imperceptible pressure tactics. The suggested mini cases can be used for teaching ethical issues in financial gerontology. Selected questions for case discussions are available from the author upon request.

Scenario 1:

An elderly woman who was living with daughter and her son in law found out that they had been spending with her credit card without her knowledge. Since she did not want to press charge against her own daughter, she then found herself liable for a substantial amount of debt. With the help of her son, she moved to an assisted-living residence. Her son also took the responsibility of managing her financial affairs. Over time she realized that the living conditions at this residence were barely endurable. When she expressed her desire to relocate into a better residence, her son threatened to stop taking care of her personal affairs including all financial matters. At ninety years of age, she was very hurt by this and she did not know where to turn to. Luckily, a niece accepted to help her and eventually she moved to better facilities and she managed to pay off her credit card balance.

Scenario 2:

Maggie, an elderly woman, was living in a low-income apartment close to her bank. She paid visit to the bank many times a week to update her accounts. This was probably an important part of her social activities. People had witnessed that she had been pressured by at least one financial advisor at that bank to invest in risky mutual funds. This financial advisor tried to convince her that she could make much more money from a mutual fund investment than from her current savings account. However, he never mentioned the higher risk associated with the funds. More importantly, he used a very authoritative tone, making her feel that she was almost guilty of doing something wrong. Given her rather timid personality, Maggie almost succumbed to the temptation. Luckily, a friend of her who happened to be at the bank could not help overhearing Maggie's conversation with the bank employee. She advised her to talk to her accountant before making a decision. Fortunately, Maggie followed her friend's suggestion.

INFORMATION/KNOWLEDGE AS AN ASSET

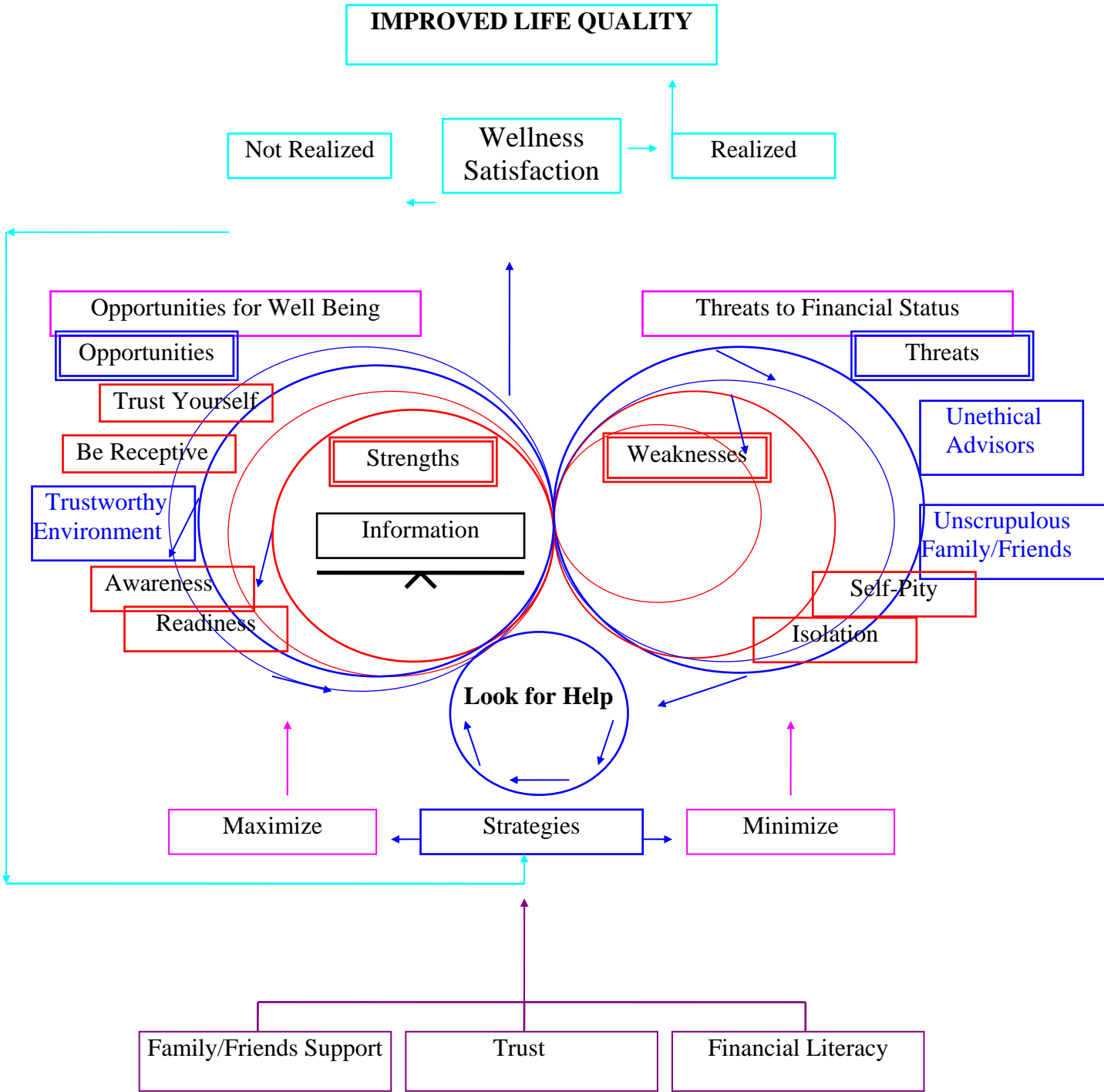
Annex 2 presents a perception of information/knowledge as an asset. As suggested by Chiappelli, Koepke, & Cherry [2], important motivators for planning are control and independence. The ongoing awareness/readiness process imbedded in the framework suggests additional motivators: higher resistance to abuses, better control of financial

assets, increased sense of ownership and safety, improved overall wellness, and avidity for more information.

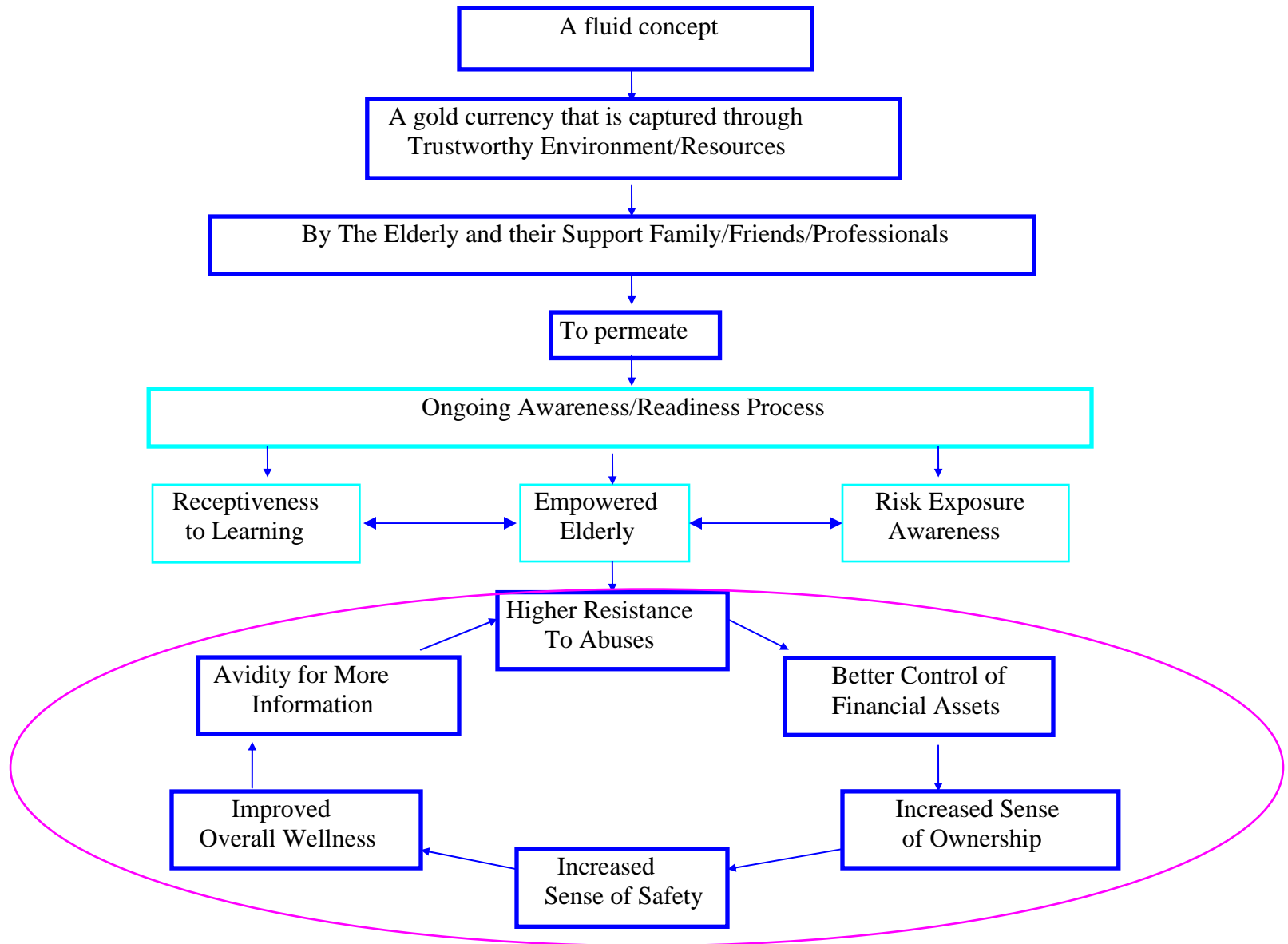
CONCLUSION

The frameworks and scenarios we propose in this paper can be used for didactic purposes. Their flexibility makes them potential workable tools for creative thinking in strategic analysis and development in financial gerontology. They can be applied for awareness improvement and thought-provoking among the elderly and concerned individuals. Future research could benefit from linking the SOWT model to quantitative and qualitative measuring tools.

ANNEX 1 - The SOWT model applied



ANNEX 2 - INFORMATION / KNOWLEDGE AS AN ASSET



REFERENCES

- [1] The American Institute of Financial Gerontology (AIFG) –
<http://www.aifg.org/aifg/aboutaifg.cfm>
- [2] Chiappelli, T., Koepke and C., Cherry K. “Planning for Long-Term Care Involves More Than Money.” *Journal of Financial Service Professionals*, 2005, 59, (2), 60-64.
- [3] Grandmont-Gariboldi N. “A Proposed SOWT Model for Teaching Strategic Analysis and Development” – *American Society of Business and Behavioral Sciences [ASBBS] E-Journal*, 2005,1(1).
- [4] Hollander A.S., Denna E.L., & Cherrington J.O. *Accounting Information Technology and Business Solutions*, McGraw-Hill Higher Education, (2000).

TROUBLING TURNOVER: AN INVESTIGATION OF DISCONTINUED E-COMMERCE MASTER'S DEGREE PROGRAMS

Marcelline Fusilier, Northwestern State University of Louisiana, (318) 357-5264,
fusilier@nsula.edu

Subhash Durlabhji, Northwestern State University of Louisiana, (318) 357-5692,
durlabhji@nsula.edu

ABSTRACT

The present study investigated characteristics of e-commerce master's degree programs that were discontinued between 2003 and 2007. Data were collected from university web sites. Of 157 total degree programs, 67 were discontinued. Findings suggested that programs offering a concentration in e-commerce were less likely to be discontinued. Business schools that were not AACSB accredited appeared to have a higher percentage of discontinued degree programs than those that were accredited. Analysis of curriculum content did not reveal discernable differences between the current and discontinued programs. This suggests that continuation decisions may be driven by factors other than program content considerations.

Keywords: Curricula, Computer applications, E-commerce.

INTRODUCTION

E-commerce has expanded since the dot-com crash. The U.S. Census Bureau News [1] reported that third quarter 2007 e-commerce retail sales increased 19.3% from the third quarter 2006. Since its inception in 1997, the annualized total return for the Dow Jones Internet Commerce Index is 21.8%. "Web Boom 2.0" is the label given by Time Magazine to the current expansion of dot-com business [2].

Consistent with the growth trend of e-commerce is the increase in e-commerce education offerings. Evidence indicates that the number of e-commerce degree programs has increased in North America [3] [4]. Various approaches to e-commerce education have been suggested in the literature: Bradbard et al., [5] advocated integrating e-commerce content into courses in traditional business disciplines. Other authors suggested taking into account the multidisciplinary nature of e-commerce and balancing technical and non-technical course material [6]. Separate e-commerce degree programs seem justified to the extent that e-commerce is a new and different way of doing business. Otherwise, a more appropriate strategy might be the integration of e-commerce content into traditional functional area business courses. At the present time, ten years after the founding of the Dow Jones Internet Composite Index, enough e-commerce experience may have accumulated to begin to address this question.

Purpose of the Study

The state of e-commerce master's degree programs in 2007 was investigated with a focus on those that have recently been discontinued. Specifically, the present study built on previous research to determine the numbers of master's programs in North America between 2003 and 2007 that have been launched, discontinued, revised, or kept the same. Characteristics of the discontinued programs were compared to those that are currently operating. The purpose is to identify potential curricular or institutional characteristics associated with program discontinuation.

Investigation of the characteristics of discontinued e-commerce degree programs may provide better understanding of factors that contribute to program life span. Findings may assist in program development and revision. These results may also contribute to future decisions concerning potential new degree programs in other areas. Given the dynamism of the business environment, there is potential for the development of new degree programs in a variety of areas. General lessons from experiences in e-commerce education might apply to decisions about potential future degree programs in other areas.

METHOD

Web Searches

The data were collected from April to July, 2007. Exhaustive web searches were conducted to find web page descriptions of e-commerce master's-level degree programs by institutions of higher education in North America. A variety of search and meta-search engines were employed in the web searches: Google, LookSmart, Lycos, MSN, AskJeeves, Dogpile, Netscape Search, About.com, Snap.com, and GoTo.com. Search terms included the program title phrases mentioned above linked with the word "degree" or "education" or "program" as well as additional qualifiers of graduate or masters.

Programs were considered e-business if their titles and/or the degree awarded contained the words e- (or electronic, Internet, or network) business or commerce, or any e-functional commerce area, such as e-marketing. Some programs that did not have any of these terms in their titles were included in the study if the curricula suggested that they were in fact e-commerce programs. Some established e-commerce programs did not have a clear web presence or one that the researchers were able to locate. In such cases, the schools were e-mailed or phoned in an effort to obtain information on their curriculum. Programs were included in the study only when a detailed curriculum description was available. Canadian web sites in French were translated by the investigators.

The degree programs took the form of (a) e-commerce concentrations in master's programs or (b) master's degrees in e-commerce. Programs resulting in an e-commerce degree require more coursework in the major area than do e-commerce concentrations. The programs were related to data collected in 2003 [3] to determine whether they were launched or discontinued during that time period. In 2003, there were 107 active programs. Continued programs' curricula were compared across 2003 and 2007 to assess whether they had been revised or remained the same. All programs were then categorized as (a) new, (b) discontinued, (c) revised, or (d) the same

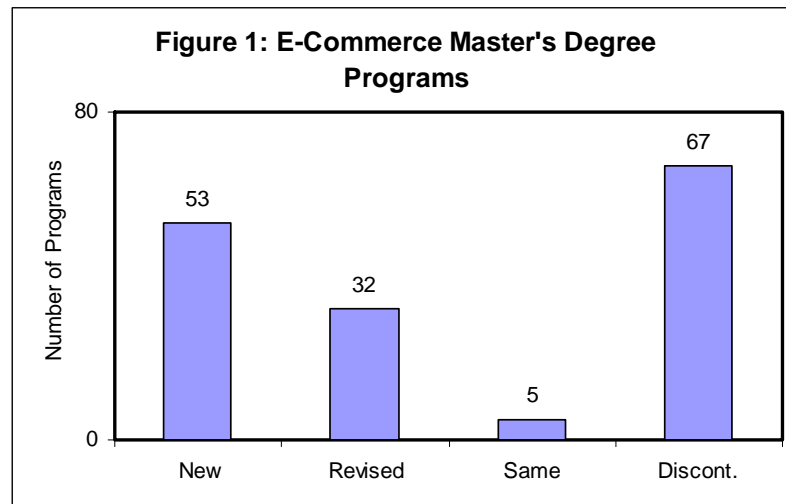
from 2003 to 2007. The college out of which the degree was offered was noted, e.g. business, computer science, etc. If the degree was offered by a college of business, that college's accreditation, or lack thereof, with regard to the Association to Advance Collegiate Schools of Business (AACSB) was recorded.

Course Coding

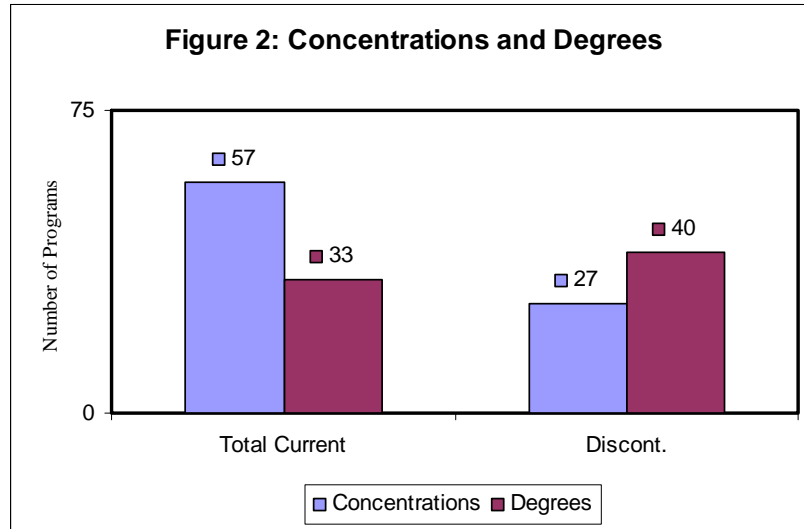
Programs were examined according to (a) number of technical courses included in the program, (b) number of e-commerce courses, and (c) number of business courses. The e-commerce courses were further identified as being technical or non-technical in nature. Courses were also classified as required or elective. Operative programs were coded according to their 2007 curricula. Discontinued programs were coded according to their 2003 curricula because they no longer existed in 2007. Course titles and descriptions were used to place courses in the following categories: (1) business, (2) e-commerce - technical, (3) e-commerce - non-technical, (4) technical, or (5) lab.

RESULTS

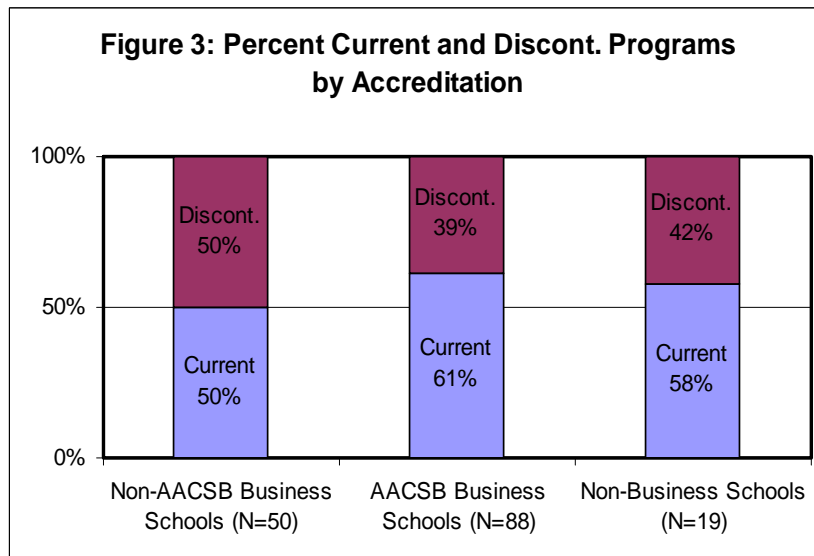
Data collection resulted in a total of 157 e-commerce master's degree programs. Of this total, 53 programs were new, 32 were revised, five remained the same, and 67 were discontinued over the 2003-2007 period. The bar chart in Figure 1 displays the program outcomes by their frequency.



Of the 90 programs actually operative in 2007, 57 (63%) were concentrations in e-commerce and 33 (37%) were e-commerce degrees. For the 67 discontinued programs, 27 (40%) were concentrations and 40 (60%) were degrees. See Figure 2. A z-test was conducted to compare the proportion of concentrations across the current and discontinued programs. Results suggested that the difference was significant ($z = 2.63$; $p < .01$).



Business schools not accredited by the AACSB offered 50 of the total 157 degree programs. Eighty-eight of the degree programs were offered out of AACSB accredited business schools and 19 degrees were housed in non-business university units, such as engineering schools. The percentages of current and discontinued programs in each group are presented in Figure 3. A z-test was conducted to compare the difference of the proportions of discontinued programs between business schools that were accredited (.39) versus not accredited (.50). Results suggested a significant difference ($z = 2.11, p < .05$).



Program outcomes (discontinued, new, revised, and same) were compared across the coding categories: business, e-commerce - technical, e-commerce - non-technical, technical, or lab. No significant differences were detected across the current and discontinued programs with regard to the average number of courses in each category.

DISCUSSION

Burkey [7] argued that electronic commerce programs are still in their infancy. If that is the case, then it has been a rough childhood. Discontinuation appears to be a pronounced trend for e-commerce master's degree programs. Durlabhji and Fusilier [3] reported that 23 master's degree programs were discontinued from 2001-2003. The present study found 67 discontinuations from 2003-2007. This is a 291% increase over the earlier period. More programs were discontinued from 2003-2007 than were launched, revised, or kept the same. This is the first time since the start of e-commerce degree offerings that the number of discontinued programs has exceeded the number of new programs [3].

Consistent with previous literature, concentrations (or tracks or specializations) comprised the majority of the currently operating programs. The prevalence of concentrations might mean that business schools now regard e-commerce more as an aspect of general business practice than as a distinctly different way of doing business and are therefore infusing e-commerce content across the curriculum. However evidence suggests that e-business has not been adequately integrated into business education [8]. Schools may be attracted to concentrations because they typically involve fewer e-business courses than degree programs and therefore tend to be more flexible and economical in terms of university resources and time demands on students.

AACSB accreditation appears associated with program continuation. The process of maintaining accreditation may require faculty and administrators to take a more measured and comprehensive approach to developing and monitoring new degree programs. This preparation and assessment may prevent or address problems before they reach a crisis stage requiring program discontinuation. The role of accreditation status in e-commerce degree programs' curriculum content and continuation is an avenue for future research.

The non-significant differences concerning the course categories (business, technical, and e-commerce technical and non-technical) suggest that program content may not be a major consideration in program continuation decisions. Bradbard et al. [5] contended that program decisions were primarily market driven in pursuit of increased enrollment. As a result, curriculum content may receive inadequate attention. Celsi and Wolfenbarger [9] addressed the problem by proposing that business schools develop "Renaissance managers" who understand the convergence between IT and the business-strategic functions. This educational approach may reflect the direction of e-commerce noted by Zhu et al. [10] for the U.S. as a technologically developed country: competitive advantage is attained not by simply adopting e-commerce, but rather by extensively integrating systems and applications based on the Internet platform. Ten years after the founding of the Dow Jones Internet Composite Index, e-commerce appears here to stay. E-commerce degree programs, accordingly, should be designed for long-term benefit.

REFERENCES

- [1] U.S. Census Bureau News, *Quarterly Retail e-Commerce Sales 3rd Quarter 2007*, November 19. Available online at: <http://www.census.gov/mrts/www/data/html/07Q3.html>
- [2] Quittner, J. "Web Boom 2.0," *Time*, December 25, 2006-January 1, 2007, 75.
- [3] Durlabhji, S., Fusilier, M. "E-business Education in Transition," *Journal of Internet Commerce*, 2005, 4(1), 103-122.
- [4] Mechitov, A.I., Moshkovich, H., Olson, D.L. "The Master's Degrees in e-commerce: A Survey Study," *Journal of Computer Information Systems*, 2002, 42(4), 29-34.
- [5] Bradbard, D.A., Robbins, D.K., Riddle, E.J., Grigsby, W.W., Stuart, E.W. "Business School Curricular Responses to the Emergence of Electronic Commerce," *Journal of Internet Commerce*, 2002, 1(1), 47-64.
- [6] Ngai, E.W.T., Gunasekaran, A., Harris, A.L. "The Maturing of e-commerce Education in our Curricula," *Journal of Information Systems Education*, 2005, 16(1), 5-8.
- [7] Burkey, J. "The Evolution of Electronic Commerce Education," *Journal of Education for Business*, 2007, 82(5), 276-281.
- [8] Morrison, J.L., Oladunjoye, G.T. "E-commerce Infusion into Business Education- Encompassing the Realities of an Emerging Business Model," *Journal of Education for Business*, 2002, 77(5), 290-295.
- [9] Celsi, R., Wolfenbarger, M. "Creating Renaissance Employees in an Era of Convergence between Information Technology and Business Strategy: A Proposal for Business Schools," *Journal of Education for Business*, 2001, 76(6), 308-312.
- [10] Zhu, K., Xu, S., Kraemer, K.L., Dedrick, J. "Global Convergence and Local Divergence in e-commerce: Cross-Country Analyses." In *Global e-commerce: Impacts of National Environment and Policy*, K.L. Kraemer, J. Dedrick, N.P. Melville, K. Zhu (Eds.), pp. 345-384, Cambridge, UK: Cambridge University Press, 2006.

INVESTMENT SYSTEMS AND SYSTEMIC ASSETS MODEL: RETURNS AND RANGES OF RISK ELIMINATION

Victor Lux Tonn, Salve Regina University, jpmapple@aol.com, (401) 848-5821
Hsi C. Li, Bryant University, hli@bryant.edu, (401) 232-6079

ABSTRACT

This paper presents an iso-xystemic study between the culture-technology expander and the virtual-real expander. For this purpose, the basic concepts of elements, impactors, and systems are briefly introduced. In addition, for the system of investments, the principle of independence will be illustrated and the principles of mutuality and centrality will be applied. In this connection, a xystematic approach to investments is also presented.

Keywords: xystematics, iso-xystemism, expander, virtual sector, investment systems

I. INTRODUCTION

In light of xystematics (systematics) -- structural, relational, qualitative, and systemic studies of things and events universal, this paper presents the mechanism of iso-xystemism (i.e., sharing essentially the same structures) between the culture-technology expander and the virtual-real expander. Consequently, a new approach to investment in terms of investment systems and virtual-real interactions is essentially presented in this paper. For this paper to be self-contained, in the Appendix, basic concepts of elements, impactors, and systems are briefly introduced, and the impacts of the virtual-financial sector on the real sector will also be examined.

II. ISO-XYSTEMISM BETWEEN THE CULTURE-TECHNOLOGY AND VIRTUAL-REAL EXPANDORS

From the culture of symbolism (i.e., culture based on the rationality of symbolism and the tranquil life-will), symbolic-technology is created [7]. The xystem (i.e., system) of symbolic culture comprises the following characteristics. (A) In terms of rationality: (1) Using the shengsheng generatics (shengsheng dynamics) as the starting point. (2) Idea of the whole (or holistic view). (3) The pattern of searching for the commonality first and then looking for the differences. (4) Thinking pattern of the unity with two dimensions. (5) Worlds of You and Wu, with yin-yang complementarity. (6) Having great concerns about the present life and the eternity of life in terms of one's offspring. (7) Things and events are thought to be maintained in coordination and complementarity. (8) Using the theoretical framework of Tcentral and the pursuit of strategic continuity or discontinuity. (9) Ideas of Tao reversal and inverse. (10) To pursue eternity in tranquility in time. (B) In terms of life-will: (1) focusing on the present life, (2) pursuit of tranquility, harmony, and peace.

From the culture of sciences (i.e., culture based on the rationality of sciences and the conquest life-will), science-based technology is created. The xystem of scientific culture comprises the following characteristics. (A) In terms of rationality: (1) Using statics as the starting point of

reasoning. (2) Tendency toward deconstructing the whole into parts. (3) Tendency toward differentiation first and then having a holistic view. (4) Dualism with antagonism in-between two forces (5) An approach of idealism. (6) Distinguishing the present world from the next life. (7) Ideas of independence and competition. (8) Using analytical framework of equilibrium and disequilibrium. (9) Employing the framework of positive existence. (10) Based on an instantaneous time to pursue eternity. (B) In terms of life-will: (1) differentiate this life from the next life, (2) the spirit of competition.

Through the intermediation of the education system, human creativities of the cultural system (of rationalities and life-wills) are carried over into the technological system. From the cultural system, the creative forces are transformed into the creative forces of the technology system. This transformation is a magnifying process, enlarging from the energy of the original cultural creativities to the one combined with the energy of the technological creativities. This is also a process that, by the interactions of the multitude of rationalities and wills of people in the communities, lower and simple levels of energies are converted into higher and complex levels of energies, and various micro energies are transformed into macro energies as well. Here, (1) the transformation of the operating forces from the cultural system into the education system is a magnifying process that raises the total level of energy, (2) the transformation of the operating forces from the educational system into the technological system is another magnifying process that leads to a higher level of energy, (3) the transformation of the operating forces from the technological system of R&D into the economic system of production and consumption is an even larger magnifying process of generating a much higher level of energy. In this last process of transforming the technological system into the economic system, the cultural creativity of the human being is combined with material's creativity in contributing to the total level of newly generated energy.

The afore-mentioned magnifying motion of the operating forces, as a mechanism of expander in inter-systemic transformations, is part of the systemic dynamics. As a culture spreading from one race or area into another, the dynamism of the old cultural forces may conjoint the dynamism of the new cultural forces. Thus, this illustrates the magnifying process of one culture interacts with another culture. For an example, the astronomy, algebra, and other pre-sciences of ancient Babylon and the geometry and still other pre-sciences of ancient Egypt were absorbed by ancient Greece and then turned into the formal seed of pre-sciences. For another example, the dissemination of the symbolic technologies of ancient China to the West contributes significantly to the original development of modern sciences in the West.

Let the mechanism of expandoor from the culture system to the technology system be iso-systemic to the mechanism of expandoor from the virtual (economic culture and finance) system to the real (production) system of the general cultural economy. In the above, by the iso-systemism of two expandors is meant that each mechanism of expandoor possesses the same structural subsystems of balance, mutuality, and centrality; furthermore, the forces of operations of all the subsystem of these expandors are perceived as essentially and respectively the same in both systems. Here, for the culture-to-technology expandoor, in the case of symbolic culture, the force of centrality is the "pursuit of things and events in totality," the force of mutuality is "to maintain things and events in proper coordination and complementarity," and the force of balance is for "all involved to be well taken care of." As for this expandoor in the case of

scientific culture, the force of centrality exhibits in the “pursuit of deconstructing the whole into parts,” the force of mutuality exhibits in “seeking independence and competition” (i.e., seeking to opposite a state of coordination and complementarity), and the force of balance exhibits in “well taking care of all involved.”

The principles of operating the symbolic financial system may be perceived as (1) seeking the commonality first and searching for differences later, (2) a process of reversal and inverse: strategic continuity and discontinuity, and (3) achieving a balance between the real and virtual sectors. The principles of operating the scientific financial system may be perceived as (1) differentiation first and seeking the commonality later, (2) a process of affirmation and negation: equilibrium and disequilibrium, and (3) the dominance of the real sector (of production) by the virtual sector (of finance).

III. INVESTMENT SYSTEMS

There are two types of investor-systems, the scientific investor-system and the double-cultural investor-system. Here, the double-cultural investor-system is a grand system that employs both symbolic culture and scientific culture in engaging in investment activities that cover both real and virtual sectors of the general economy so that the investment system is in a state of balance with regard to its direction of investments toward the real sector and the virtual sector (or financial sector). This is because of its cultural inclination toward the composite (grand) system comprising a symbolic financial subsystem and a scientific financial subsystem.

Scientific investor-system is a system in pursuit of profit maximizing opportunities of investment activities. This system does not concern itself with the operations of the real production sector of the economy for the sake of real production. In contrast, the double-cultural investor-system would consider the real production activity as a desirable feature closely associated with its assets of investments. This guiding principle of the scientific investor-system for profit maximization is consistent with its underlying cultural orientation toward the idea of employing the virtual forces in controlling the real forces of the world surrounding us.

The decision-making “star-7 principle”(i.e., low level unified jT principle [7]) for investors is the principle with the following ranking order: holistic advancement, centrality, mutuality, balance, independence-creativity, aesthetic beauty, and transcendental concerns. This principle also comprises the least acceptable and least satisfactory principle, dissatisfactory but accepting principle, constantly changing status quo principle, trial and error principle, non-profit principle, optimization, and others in striving for achieving the minimally acceptable state of satisfaction/dissatisfaction or the lowest acceptable level of disequilibrium.

Practically, in the virtual-financial sector, on the demand side, there are the following investment systems: benchmark-investors, lazy investors, familiarity-investing system (those investing in the corporations with which the investors are familiar), day traders, and long-term investors, etc. On the supply-systems, there are those companies (i.e., systems) issuing primary securities and those offering secondary and other securities.

In the real sector -- the real grand system -- of the grand cultural economy, there are three subsystems: subsystem of competition, humanity/socio-academic subsystem, and government subsystem, with each having its own control mechanism ([7]). Let the money-financial sector -- the virtual sector in the narrow sense -- of the grand cultural economy, be a grand system having three subsystems: subsystem of competition, humanity/socio-academic subsystem, and government subsystem, each having its own control mechanism as well. Thus we meant, there is an *iso-xystemism* (i.e., the relationship of sharing *fundamentally* the same structures in the sense of xystematics) between the virtual (or financial) sector and the real (or production and consumption) sector of the grand cultural economy.

In the subsystem of competition in the virtual-financial sector of the grand cultural economy, the sub-subsystem of balance and competition is defined as a system with the collection of all forces of balance and competition in this subsystem. The α -side factors of this subsystem of competition are the investors including the individual investors and the institutional investors, and the β -side factors are the private firms that issue stocks (or financial assets).

In the same subsystem of competition in the virtual-financial sector, the sub-subsystem of mutuality comprises the forces of mutuality and coordination. The α -side factors of this sub-subsystem are the interactive and friendly forces among the individual investors (such as in investment clubs) and the friendly and non-profit-oriented interactive forces among the investment institutions (also see [3], pp.10-13). As for the β -side factors, they are the forces of following-up after significant price changes.

In the same subsystem of competition, the sub-subsystem of centrality comprises the forces of centrality. The α -side factors of this sub-subsystem are the Fed, the roles of leadership among the individual investors and the roles of leadership among the institutional investors (also see [1][2]). As for the β -side factors, they are the forces responsible for artificially generated big price changes.

We reason that there is an iso-xystemism between the culture-to-technology expander and the investors-to-virtual (system) expander. That is, we mean that the culture-to-technology expander has structures and forces of centrality, mutuality, and balance, and the investors-to-virtual expander has the structures and forces of centrality, mutuality, and balance as well.

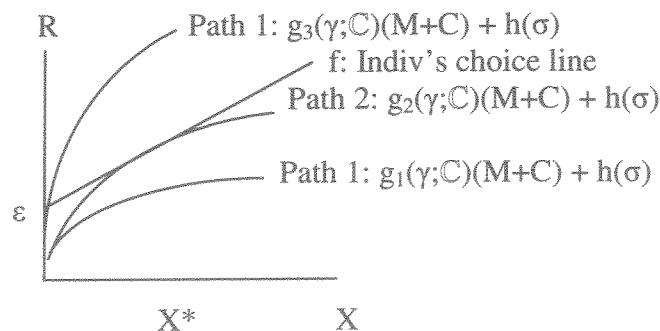


Figure 3.1 Determining the level of investments

Now we introduce the investment decision model (IDM). Let A be a system of our interests and g be the growth path of the forces underlying the trend growth of system A. Let R denote the expected asset returns (using e.g., symbolic rationality) of the system A, f be the line of the individual investor's expected gains from investments (emanating from the intercept ε on the vertical axis R), $g_i(\gamma)$ (with γ being the growth factor, i representing various scenarios from bad to good states) be the factor growth path of portfolios, with the curve derived by arranging portfolios from low to high levels of expected returns, and $h(\sigma)$ be the risk factor (σ being the standard deviation of R , as in the CAPM model [4], pp. 148-173). In Figure 3.1 (also see Figure 3.2), we draw three factor growth paths (emanating from the origin) representing three scenarios of trend growth. Then we have:

$$(3.1) \quad \text{Return: } R = g_i(\gamma; C)(M+C) + h(\sigma),$$

with C being the forces of centrality, M being the force of mutuality, and C being the cultural variables (comprising the culture in general, institutional and organizational factors, and strategic factor) as is in the case of a cultural growth model [7].

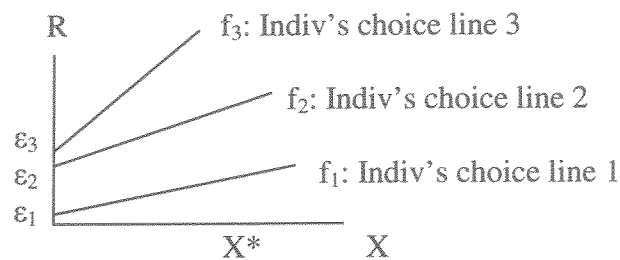


Figure 3.2 Higher ε_i leads to higher slope

IV. BRIEF CONCLUSION

This paper presented an iso-systemic study between the culture-technology expander and the virtual-real expander. For the systems of investments, the principles of competition, mutuality, and centrality were respectively applied.

APPENDIX SYSTEMATICS AND THE VIRTUAL-REAL IMPACTS

By an element, represented by a letter (i.e., an alphabet in any spelling language) or any other symbol (such as a Chinese character), is meant an entity expressing the idea of a thing existing in the universe, be it physical, cultural, or even imaginary. Primitive element is an undefined term, while, in general, a composite element is an organic and integrated entity composed of primitive elements -- simple element(s) or impator(s). The most basic attributes of human being and the characteristics of natural entities are also primitive elements. In general, a material element is an organic entity (i.e., a composite element) composed of most basic universal elements, universal elements that denote physical characteristics, and some impactors. In general, human element is an organic entity (i.e., a composite element) composed of universal elements denoting physical attributes, human characteristics, and some impactors. The impactor refers to the force of transformation of one element, xystem or else into other. The idea of xystemism (i.e.,

transsystem) refers to the transformation of one thing -- element, xsystem, thing, event, or entity - into another [7]. Similarly, we have the ideas of systemisms [5] [6].

A system (i.e., system) is defined as an organic composition of elements ($x_i^\#$ in $\langle x_i^\# \models \gamma \rangle$, γ denoting the group culture) having a set of static and dynamic structures (Ψ), a set of the group/social control mechanism (“ τ ” in $Z \equiv Z(\tau)$), a set of decision-action and strategic set of elements (Z), and a set of cultural complex ($\{C\}$) [5] [6][7]. Formally, for a system X , we write: $X = \langle x_i^\# \models \gamma \rangle^{(L)} \langle \Psi-Z \parallel E \rangle \oplus \langle C_X \rangle$, with $\langle \Psi-Z \parallel E \rangle$ referring to the complex of static-dynamic structures of social relationships and actions-strategies conditioned upon the external environments (E), with “ (L) ” depicting the uncertain forces; “ $\langle \rangle$ ” representing the notations of the constituting elements with their group cultures, “ $\langle \rangle$ ” representing the complex of the static-dynamic structures and others, and “ \oplus ” representing an inter-connecting impactor. The space of systems is then the collection of all systems joined by the collection of operating forces (called impactors) and the family of cultural spaces.

For an economy of triple control, the favorable areas of investment areas for the government subsystem are: relatively large scale projects, industries of long-term digestion and with matured technologies, military industries, industries of advanced and not-yet-matured technologies, and other projects involving the national prestige. The favorable areas of investment areas for the investors in the humanity/social-academic subsystem are: cultural products and those industries with relatively new and high-level technologies. The favorable areas of investment areas for the investors in the competitive subsystem are: real estates, areas related closely with the daily living of the common people, industries with matured technologies, areas involving relatively small-scale capital requirements, short-term digestion period, and flexible operations.

The principles of investment operations are as follows. The principle of centrality is: expected-profit maximization. The principle of balance/mutuality is: diversification of investment areas. The principle of independence is: to be different by taking on a surprise factor.

REFERENCES

- [1] Alfarano, S. Lux, T. “A Simple Model of Herd Behavior with Realistic Time Series Properties,” in A. Kirman A. Teyssiere, G. eds., *Long Memory in Economics and Econometrics*. Berlin: Springer, 2004.
- [2] Drehmann, M. Oeschssler, J. Roider, V. Herding with and Without Payoff Externalities – An Internet Experiment, 2006.
- [3] Nofsinger, J. R. *The Psychology of Investing*. Upper Saddle River: Pearson Education, 2002.
- [4] Pyndyck, R. S. Rubinfeld, D. L. *Microeconomics*, 5th Ed. Upper Saddle River, N.J.: Prentice-Hall, 2001.
- [5] Tonn, V. L. “Lecture Notes on The Economics of Culture” (mimeo.), 2001.
- [6] Tonn, V. L. *Systematics and The Economics of Culture*. Frankfurt am Main: Peter Lang 2008 (forthcoming).
- [7] Tonn, V. L. Culture and Systemic Competition/Cooperation between Nations: global economic framework with a model of catching-up process (mimeo.), 2008.

INTEGRATING SERVICE-ORIENTED ARCHITECTURE (SOA) AND COMPUTING (SOC) METHODOLOGY INTO INFORMATION SYSTEMS CURRICULA WORKSHOP

H. Howell-Barber and James Lawler, Pace University, 163 William Street, New York, New York 10038, 212-346-1013, h.howell@verizon.net and jlawler@pace.edu

PARTICIPANTS

H. Howell-Barber, Pace University

James Lawler, Pace University

SESSION FORMAT

Service-Oriented Architecture (SOA) is defined as an enabling framework for improving business processes for competitive advantage. The complexity of SOA and the contrast of SOA to elementary web services create however a challenge in attempting to define an approach to deployment in firms and to education. This workshop defines a practical program management methodology to address the complexity and the hype of SOA as a technology and to integrate into information systems curricula.

The methodology of SOA is defined in frameworks of best practices for participants on projects of SOA. They distinguish the program management methodology from established project management methodology in information systems curricula and furnish principles of service orientation that may be integrated in the curricula. The workshop describes essentially an evolutionary methodology of non-agile and agile project management techniques that are distinguished by the participants in the customization of the techniques to SOA.

This workshop will benefit instructors attempting to educate students on the complexity of the technologies behind SOA. The workshop will concentrate on the business dimensions of SOA as more critical than the functionality of platform technologies and will cover case studies of firms in industry that have deployed and expanded web services to SOA, integrated process and services architecture, and exploited enterprise services, based on a service-oriented enterprise (SOE). This workshop will include a foundation for a syllabus on SOA that can be integrated into information system curricula for undergraduate and graduate students.

AN APPROACH TO IDENTIFY UNDERVALUED COMPANIES

Raja R. Vatti
St. John's University
Jamaica, NY 11439
718/990-6338
vattir@stjohns.edu

ABSTRACT

In the process of portfolio selection, investment analysts take mainly two approaches. One is a technical approach which concentrates on historical patterns in the price behavior of a stock. The other is a fundamental evaluation of the historical financial performance and the future prospects of a company. Many optimization methods and statistical techniques were used in the past in order to select portfolios with good positive return potential. The current investigation explores the feasibility of applying a multivariate statistical segmentation technique to identify one or more undervalued companies in each chosen industry. The procedure looks at some key financial performance variables of the previous year for a sample of companies in an industry and classifies companies into groups that are highly similar in fundamentals. The companies with decent reward potential are then identified.

INTRODUCTION

In the fundamental analysis, investment analysts analyze all the information available on the economy, industries, and companies in order to identify the companies with good price appreciation potential. Some make decisions subjectively, and there are others who base their decisions on a detailed analysis of the data.

In the past forty-five years, there was significant research to establish the value of a company's reported financial information in the investment analysis. Ball and Brown in 1968 were among the first who investigated the value of the accounting data in explaining the security prices. In the Accounting Review of July 1986, Senyo Tse also demonstrated the extent of the relationship and the changing trends in the coefficient of determination using some cross-sectional models with approximately 420 firms in the sample.

There were many other regression models that utilized financial information in establishing Beta values, systematic risks, yield spreads, earnings etc. In addition to the regression analysis, there were many applications of the other multivariate statistical analysis.

Discriminant analysis was used in identifying troubled companies and there were factor analytic models for establishing potential stock returns. Several time series models were also developed to forecast earnings and prices.

The accounting numbers and ratios that were frequently analyzed dealt with a company's sales, earnings, dividends, debts, cash flows, equity, liquidity and assets.

CURRENT RESEARCH

The objective of the current study is to investigate the applicability of the multivariate classification techniques in the context of portfolio analysis. Traditional belief is that all companies within an industry tend to behave similarly for the changes in the economic environment. It is generally assumed that the observed differences among the companies of an industry are due to the operating efficiencies or deficiencies.

By taking a sample of companies in the educational services industry, we want to investigate whether further segmentation could be performed within the industry using some relevant financial performance ratios calculated from the reported financial statements. Groups or clusters of companies can be formed by analyzing the latest information such as earnings yield, cash flow yield, book to price ratio, dividend yield, debt to market price, working capital to price and net income margin.

The companies that are closest in the sense of

Squared Euclidean distance with respect to these financial variables will be put together in a cluster. At the end of the classification procedure, an attempt will be made to determine whether the price behavior of the companies in a cluster is similar. It is anticipated that the companies in different clusters would have different levels of price changes.

James Farrel, in the May/June 1975 Financial Analysts Journal, published one similar application of the cluster analysis. He analyzes monthly price changes of several companies in several industries, and he forms clusters based on the similar price behavior during 1961-1969.

The current analysis differs from the analysis of Farrel's since his analysis is restricted to the price behavior over several years. The plan here is to analyze several measures in a given year, and we want to explore the feasibility and practicality of observing similar companies in the portfolio analysis. One or more undervalued companies in very high similarity group or groups will then be identified.

DATA

Financial performance data for the year 2006 was taken for the Educational Services industry from the Value Line Investment Survey reports. For each company selected in the chosen industry, the variables included in the data are

- Revenue yield
- Earnings yield
- Cash flow yield
- Working Capital yield
- Book value yield
- Net income margin
- Change in Revenue yield from the previous year
- Change in earnings yield from the previous year

ANALYSIS

Using the performance variables for the chosen companies in the Educational Services industry, the companies are classified into several similar groups using squared Euclidean distance as the criterion for the closeness of the companies. The procedure used is the Wards' hierarchical clustering. This method would minimize the variation within the groups. The

similarity measure for each group is noted, and one or more undervalued companies are selected from the groups with a very high similarity. Actual price appreciation of the identified stocks during the next year would be compared against the other companies in the same group

RESULTS

Table 1

<u>Group</u>	<u>Similarity Level</u>	<u>Companies Formed into Group</u>
1	97.82	Apollo group, ITT Educational Services
2	97.34	Renaissance Learning Strayer Education
3	96.47	Career Education, Devry
4	N/A	Bright Horizons
5	N/A	Corinthian Colleges
6	N/A	Skillsoft

If we take the top similarity group with similarity level of 97.82, there are two companies in it with close fundamentals. Between the two, ITT Educational Services has higher earnings and revenue yield and higher return on equity.

In the second top group with similarity level of 97.34, Strayer Education appears to be undervalued with its higher earnings yield and return on equity.

From the two groups, ITT Educational Services and Strayer Education appear to be undervalued companies.

Table 2

<u>Group</u>	<u>Company</u>	<u>Actual 2007 % Stock Price Appreciation</u>
1	Apollo ITT Educational Services	22.8 30.8
2	Renaissance Learning Strayer Education	-9.6 65.1

CONCLUSIONS AND FUTURE RESEARCH

The proposed multivariate statistical procedure could successfully identify two undervalued companies in the Educational Services industry.

If the two identified companies were included at the beginning of 2007 in an investment portfolio, the returns from the two would have been higher than the returns from the other companies in the groups.

The method should be tried in several industries to establish it as a viable tool for portfolio selection.

The Ward's classification method should be compared against the other clustering mechanisms to select the most appropriate one.

REFERENCES

1. Anderberg, M.R. (1973), Cluster Analysis for Applications, Academic Press.
2. R. Ball and P. Brown. "An Empirical Evaluation of Accounting Income Numbers,"
Journal of Accounting and Research, Autumn 1968, pp. 159-178
3. Farrel, J. "Homogenous Stock Groupings",
Financial Analysts Journal, May-June 1975, pp.50-62
4. Tse, S. "Intra-Year Trends on the Degree of Association between Accounting Numbers and Security Prices", Accounting Review, July 1986, pp. 475-497

ID Theft and its Financial Implication
Hadley Leavell, Sam Houston State University
leavell@shsu.edu

ABSTRACT

This paper will address the methods that the ID theft criminals use to collect and access identity information and how they are using this information to commit theft and fraud. This paper will also address the financial impact that identity theft has on consumers, businesses, and governments. The laws that are currently in effect relating to ID theft will be covered and new ideas for the prevention of identity theft and the safeguarding of identity information will be addressed.

THE SUBPRIME MORTGAGE CRISIS: THE IMPACT ON FINANCIAL MARKETS

Luis Eduardo Rivera-Solis, Dowling College, riveral@dowling.edu

ABSTRACT

During the period 1998 through 2006 millions of subprime mortgages were provided by financial institutions. There has been a high foreclosure rate despite the fact that interest rates have been low and the economic environment has been favorable. It is because of these high foreclosure rates, many financial institutions have experienced losses and it has caused concern in the financial markets. The purpose of this paper is to explore subprime mortgage crisis and its impact on the financial markets.

REFERENCES

- [1] Danis, Michelle and Pennington-Cross, A., "The Delinquency of Subprime Mortgages", Working Paper 2005-022A, <http://research.stlouisfed.org/wp/2005/2005-022.pdf>.
- [2] Feldman, R. and Schmitt, J., "Why All Concerns About Subprime Lending Are Not Equal", *Fedgazette*, Federal Reserve Bank of Minneapolis, July 1999, <http://www.minneapolisfed.org/pubs/fedgaz/99-07/banking.cfm?js=0>
- [3] Schloemer, E., Li, W., Ernst, K. and Kees, Kathlee, K., *Losing Ground: Foreclosures in the Subprime Market and the Cost to Homeowners*, Center for Responsible Lending, December 2006., <http://www.responsiblelending.org/pdfs/foreclosure-paper-report-2-17.pdf>

INSIDER HOLDINGS AND STOCK VALUATION: A STATISTICAL ANALYSIS

Steven V. Le, Department of Finance, California State University, Long Beach
1250 Bellflower Blvd., Long Beach, CA 90840-8505, (562) 985-1608, svle@usa.net

ABSTRACT

This statistical analysis, based on monthly financial data of 1,810 Value Line companies (98 industries) between January 2003 and August 2007, reveals that high percentage of outstanding shares held by insiders (corporate Officers and Directors) has a positive impact upon the firm's intrinsic value.

I. INTRODUCTION

Business Week (2003) reported that one-third of the S&P 500 companies have founding families involved in management, and these are usually the best performers. Shleifer and Vishny (1997) observe that the large premiums associated with superior-voting shares or control rights provide evidence that controlling shareholders seek to extract private benefits from the firm. Family ownership and control, in public U.S. firms is commonly perceived as a less efficient and less profitable ownership structure than dispersed ownership. Stein (1988, 1989) shows that the presence of shareholders with long investment horizons can mitigate the incentives for myopic investment decisions by managers.

While empirical studies generally have supported the hypothesis that interest rate and dividend changes affect common stock prices [see Feldstein (1980, 1983), Hong (1977), Modigliani (1979), Matley (1976), Van Horne (1972), and robe (1969), Cottle (1962), Gordon (1962), Pettit (1977), Rosenberg (1979), and Litgenberger (1979, 1980, 1982), Bar-Yosef (1987), Kalay (1986), Ofer (1987), March (1987), Ambarish (1987), Brickley (1983), Bhattacharya (1979), Hakansson (1982), Asquith (1983), Miller (1982, 1983), Afarony (1980), Charest (1978), Pettit (1976), and Watts (1973)], no statistical study has been done to investigate the effect of percent of insider holdings on the valuation of publicly owned companies stocks.

The objective of this study is to examine the relationship between insider ownership (percentage of outstanding shares held by insiders) and valuation of large public firms. The remaining sections of this paper are organized as follows. Section II presents the statistical model, methodology and data. Section III discusses the empirical results. The conclusions are in Section IV.

II. METHODOLOGY AND DATA

The statistical model constructed for this study is based on the generally accepted theory of common stock valuation. This approach is based on the principle that rational investors evaluate the expected returns and risks of securities in the financial market and set a price for a particular security which adequately compensates investors for the risks. The Discounted Cash Flow valuation approach is based on the proposition that the maximum price that a rational investor will pay for a security is an amount equal to the present value of the expected dividends plus its resale price, including capital gains. Therefore, the present market price of a stock is given by the formula:

$$P_0 = \frac{D_1}{(1+K)^1} + \frac{D_2}{(1+K)^2} + \dots + \frac{D_t}{(1+K)^t} + \frac{P_t}{(1+K)^t} \quad (1)$$

Equation (1) was simplified by Gordon (1962) as follows:

$$P_0 = \frac{D_1}{K-g} \quad (2)$$

Where g is the expected dividend growth rate.

Equation (2) can be expressed as follows:

$$\frac{P_0}{B_0} = \frac{D_1/B_0}{K-g} = f(D_1/B_0, K, g) \quad (3)$$

Where P_0 / B_0 = market price-to-book ratio
 B_0 = book value
 D_1 / B_0 = book yield
 K = $R_f + \text{risk}$
 R_f = Risk-free rate

Equation (3) attempts to quantify the impact and the relationship between stock prices and a number of economic, financial and risk factors associated with each company. The ratio of market price and book values of security i can be written as a function of several explanatory variables and can be expressed as follows:

$$P_i/B_i = f(RF, \text{book yield}, g, \text{risk}) \quad (4)$$

There are four types of variables which were hypothesized to affect the market price-to-book ratio of companies:

- (1) Economic Variables: Interest rates and inflation should have an effect on market price-to-book ratio.
- (2) Dividend Policy: High book yield, retention ratio, and expected earnings growth rate should have a positive effect on market price-to-book ratio.
- (3) Risk Factors: high debt, high beta, and low Value Line Safety Rank should have a negative impact on market price-to-book ratio.
- (4) Financial Factors: High return on equity, high percent of cash to total asset, good Value Line Timeliness Rank, high dividend growth, positive money flow, and high annual return should have a positive impact upon market price-to-book ratio.

In specifying (4), our intent is to construct a statistical model to quantify the changes in the market price-to-book ratio and to examine the relative importance of insider holdings versus other economic and financial factors in the valuation of stock prices.

This empirical study is based on monthly Value Line financial and economic data from January 2003 through August 2007 of 1,810 companies (98 industries). The monthly data was obtained from Value Line and Federal Reserve Statistical Release. The dependent and independent variables were defined as follows:

- **Market/book ratio (P_0/B_0):** The month-end market price divided by book value per share.
- **Book yield (BYD):** Indicated declared dividend divided by book value per share.
- **Risk-free rate (I):** The interest rate of the 2-year U. S. Treasury Bonds.
- **Timeliness Rank (TR)** measures probable price performance during the next 6 to 12 months, relative to all other 1700 stocks. These 1700 equities represent 94% of the trading volume on all U.S. stock exchanges. The rank of a stock's probable relative market performance in the year ahead. It is derived by a computer program using as input the long-term price and earnings history, recent price and earnings momentum, and earnings surprise. All data are known and actual. Stocks ranked 1 (Highest) and 2 (Above Average) are likely to outpace the year-ahead market. Those ranked 4 (Below Average) and 5 (Lowest) are not expected to outperform most stocks over the next 12 months.
- **Safety Rank (SR):** A measurement of potential risk associated with individual common stocks. The Safety Rank is computed by averaging two other Value Line indexes – the Price Stability Index and the financial strength Rating. Safety Ranks range from 1 (Highest) to 5 (Lowest).

- **Beta (B):** A relative measure of the historical sensitivity of the stock's price to overall fluctuations in the New York Stock Exchange Composite Index. The Betas are adjusted for their long-term tendency to converge toward 1.00. Additionally, Value Line shows betas computed based on monthly total returns for the trailing three year, five-year and 10-year periods.
- **Relative P/E Ratio (RPE):** A stocks price-earnings ratio divided by the price-earnings ratio for a market measure.
- **% Retained to Common Equity (RR):** Net profit less all common and preferred dividends divided by common equity including intangible assets, expressed as a percentage.
- **Estimated Return on Shareholders Equity (ROE):** Indicator of profitability. Determined by dividing net income for the past 12 months by common stockholder equity (adjusted for stock splits). Result is shown as a percentage.
- **Dividend Per Share Growth 10-Year (DG)**
- **Total Return 1-Year (TT):** The capital gain or loss for the stock price plus the sum of dividends reinvested at year-end for the past year, expressed as a percentage.
- **Relative Strength 1 Week (RS):** The stock's price over time divided by the Value Line Composite Average over the same time span. Arising relative strength line means the stock has been outperforming the market; a declining line means just the opposite.
- **1-Month Money Flow (MF)**
- **Projected 3-5 Year Relative P/E (PPE)**
- **Projected Earning Per Share Growth Rate (PEG):** The estimated growth rate in earnings expressed as a percentage.
- **% Insider Holdings (IH):** The percentage of outstanding shares held by insiders (corporate officers and Directors).
- **Volatility Index (VIX)**
- **% Debt/Capital Latest Quarter (D)**

Utilizing a cross sectional time series data, this model may be expressed as follows:

$$P_{it}/B_{it} = a + b_1TR_{it} + b_2SR_{it} + b_3B_{it} + b_4RPE_{it} + b_5RR_{it} + b_6ROE_{it} + b_7DG_{it} + b_8TT_{it} + b_9RS_{it} + b_{10}MF_{it} + b_{11}PPE_{it} + b_{12}VIX_{it} + b_{13}PEG_{it} + b_{14}IH_{it} + b_{15}BYD_{it} + b_{16}I_{it} + b_{17}D_{it} + e_{it} \quad (5)$$

Where: i = company i
 t = time t
 a = the intercept
 b = regression coefficient
 e_{it} = the random error

III. EMPIRICAL RESULT

As shown in Table 2, a cross-sectional regression estimate of expression (4) and (5) yield the following result:

$$\begin{aligned}
 P/B = & -1.254 - 0.144 TR - 0.132 SR - 0.222 B + 0.936 RPE + 0.107 RR + 0.097 ROE + & (6) \\
 & (-15.312) \quad (-11.433) \quad (-6.60) \quad (33.64) \quad (61.84) \quad (48.594) \\
 & 0.003 DG + 0.004 TT + 0.002 RS + 0.001 MF + 1.150 PPE - 0.024 VIX + \\
 & (5.085) \quad (15.749) \quad (7.098) \quad (2.710) \quad (35.622) \quad (-13.555) \\
 & 0.022 PEG + 0.005 IH + 12.068 BYD - 0.004 I - 0.004 D + e_{it} \\
 & (22.671) \quad (9.230) \quad (48.032) \quad (-5.654) \quad (-8.859)
 \end{aligned}$$

(t-statistics in parentheses below the coefficients) ($R^2 = 0.86$)

Durbin-Watson *test* was utilized to test the hypothesis of no autoregression. As shown in Table 2, the Durbin-Watson statistic of 1.94 indicates that there is no autoregression and we can retain the statistical estimates without concerning a bias of the estimated standard error. The low correlation coefficients of the correlation matrix indicate little multicollinearity between the independent variables.

TABLE 2
STATISTICAL RESULTS

Dependent Variable: P/B: Market Price/Book Value			
Independent Variables	B	Standard Error	t
		B	
TR: Timeliness Rank	-.144	.009	-15.312
SR: Safety Rank	-.132	.012	-11.434
B: Beta	-.222	.034	-6.600
RPE: Relative P/E Ratio	.936	.028	33.646
RR: % Retained to Common Equity	.107	.002	61.842
ROE: Est Return on Shareholders Equity	.097	.002	48.594
DG: Dividend Per Share Growth 5-Yr	.003	.0006	5.085
TT: Total Return 1-Year	.004	.0003	15.749

RS: Relative Strength 1 Week	.0032	.0003	7.098
MF: 1-Month Money Flow	2.378	.877	2.710
PPE: Proj 3-5 Yr Relative P/E	1.150	.033	35.622
VIX: Volatility Index	-.024	.0017	48.032
PEG: Proj EPS Growth Rate	.022	.001	22.672
IH: % Insider Holdings	.0049	.0005	9.230
BYD: Dividend Declared/Book Value	12.068	0.251	48.032
I: 2-Year Treasury	-.0039	.0007	5.654
D: %Debt/Capital Latest Quarter	-.0037	.0004	-8.859
(CONSTANT)	-1.254		
R Square	.86		
Adjusted R Square	.86		
Durbin-Watson Statistic	1.94		

The empirical results indicated that high percentage of outstanding shares held by insiders (corporate officer and Director) would have a positive impact upon the company's intrinsic value.

The statistical results indicated that investors respond positively to the stocks with high dividend and quality earnings, which is reflected in the book yield and return on equity variables.

The results suggest that expected growth in earnings or capital appreciation is an investment objective of stockholders. This is consistent with the discounted cash flow approach in the valuation theory of common stock.

All of the financial risk factors are significantly related to the valuation of common stocks, the evidence suggests that low debt ratio, and good Timeliness and Safety Ranks would have a positive impact upon stock prices.

IV. CONCLUSIONS

This study examines the relationships between stock prices and insider stock holdings. The empirical results led to the following conclusions:

- The percentage of insider holdings (corporate officers and Directors) is seriously being taken into consideration by investors in evaluating public companies stocks. Companies with high percentage shares held by insiders will have a positive impact on the companies' intrinsic value.
- The empirical evidence suggest that high projected earnings growth, return on equity, quality earnings and good balance sheet would have a positive impact upon the value of common stocks.

AN INVESTIGATION OF GOLD PRICE

Raja R. Vatti
St. John's University
Jamaica, NY 11439
718/990-6338
vattir@stjohns.edu

ABSTRACT

Dollar value in Euros and Gold price in dollars have been moving in the opposite directions during the past several months. The investment analysts suspect that the consistent decline in dollar value relative to the Euro currency could be the major reason for the steady climb of the gold price. It is presumed that investors are preferring to hold gold instead of dollars, and as a consequence, there is an increased demand for gold. The World Gold Council announced in 2004 that a study with the weekly data on gold price and some individual country currencies in Europe from 1971 to 2002 revealed a strong negative correlation between a chosen country dollar exchange rate and dollar gold price. The current investigation is concerned with the quantification of relationship between the Euro exchange rate and dollar gold price using daily and weekly data from January, 2006 to October 2007. It is believed that the most recent data is more relevant for the analysis since the gold price has been experiencing a very rapid appreciation during the past two years.

INTRODUCTION

Gold price has been climbing steadily during the past few years, and there is very little published research with detailed investigation and search for its reasons. It has been widely believed for many years that holding gold is a good hedge against inflation. Even in ancient times, Gold was perceived as a valuable commodity with a lasting value giving it's holder unshakable sense of security and comfort in the times of personal, and national economic and political crises. It is possible that the special place of gold as a source of wealth is due to its unusual luster, chemical properties, scarcity, high liquidity, and marketability. Gold is held not only by individuals but also by the national governments as a backup for the currencies issued for circulation. It is also true that there is a great demand for gold jewelry in countries like India, and naturally, when there is

more demand than supply, the commodity's value and price increase.

A new source of demand for gold arose after 1971 when the major countries of the world allowed their currencies to float after the collapse of the Breton Woods system in 1971, and it offered some great new investment opportunities for the International traders. The fluctuations in major currencies increased the probabilities for large gains and losses in the currency trading. The tremendous increase in risks due to volatility in currency exchange rates made the investors look for means to reduce the inherent risks that are involved. The currency traders discovered that one way of hedging is to invest simultaneously in gold contracts.

At least until now, the dollar is perceived as the main trading currency in the world, and therefore, it appears to be reasonable to assume that the gold investments and the dollar value fluctuations could somehow be related because the traders attempt to protect themselves from the anticipated dollar value erosions by simultaneous investments in gold. Despite the logical intuitive presumption of a close association between the world's premium currency and the dollar gold price, the published research about their relationship did not materialize until the release of a report of the World Gold Council in 2004. The Gold Council released the results of a statistical research conducted by three academic professionals in the United Kingdom which might have been supported by the World Gold Council.

PREVIOUS RESEARCH

In 2004, the World Gold council released the results of a research conducted by Forrest Capic, Terrence Mills and Geoffrey Wood. Their research established for the first time the quantitative relationship between gold price and

the currency exchange rates of a few currencies against the U.S. Dollar. The end of week U.S. dollar prices of the currencies such as the German Mark, British Pound, Swiss Franc, and Japanese Yen were analyzed along with the end of the week gold prices from 1971 to 2002. They found a significant negative correlation between the U.S. Dollar Gold Price and the dollar price of foreign currencies such as the German Mark, British Pound or Japanese Yen. They attempted to show t gold as a consistent hedge against the exchange rate risks. It appears that their investigation is supported by the World Gold Council. The World Gold council is a marketing organization for the major gold mining companies. Their key finding is that the US dollar gold price moves in opposition to the US dollar price of the major currencies. Their observation was that if the US dollar appreciates, the dollar gold price declines. On the other hand, a decline in the value of US dollar relative to any major currency results in a rise in the dollar gold price. The researchers also concluded that the relationships were stable even when the total study period was broken into smaller variable time periods.

CURRENT RESEARCH

The main objective of the current research is to study the daily fluctuations in the dollar value of Euro currency and Gold price during the last couple of years in order to investigate the relationship between the US dollar price of currency and the dollar gold price. The current investigation is different from the research released by the World Gold council. The previous research considered data only up to 2002. There were dramatic changes in the U.S. Economy, global technology, politics, and gold prices after 2002, and therefore, the current investigation is interested in analyzing the data of most recent years after the year 2002. Instead of targeting major currencies of Europe as in the previous study, we target Euro currency only in the Europe. It is very relevant here to establish whether any significant changes have taken place after 2002. Moreover, the present study deals with the daily data instead of the weekly data. It would be interesting to explore the differences and similarities between the two investigations. It is possible that the World Gold council was interested in publicizing the importance of Gold as a hedge against the currency exchange risks. Their main objective was to increase the demand and interest for gold. It is also conceivable to think that some information could be lost if the basic

daily data is aggregated to weekly or monthly formats. Therefore, the purpose of the current study is to establish whether there is a significant relationship between the daily dollar value of a particular currency and the daily Gold price. If there is a relationship, an attempt would be made to identify the type of relationship. An attempt would be made to determine whether the dollar value is the dominant factor that is responsible for most of the fluctuations in the gold price. The research also addressed to differences and similarities between daily and weekly analysis. The current type of analysis has not been reported before, and the results could be very relevant for investment industry.

DATA

The daily London afternoon gold price, and the daily Euro exchange rate history from the U.S. Federal Reserve Bank releases were collected from January, 2006 to October, 2007. Weekly data and monthly data was established from the basic daily data. Every Friday's data becomes the end of week data, and the observations at the end of each month became the monthly data.

ANALYSIS

A scatter diagram between gold dollar price and value of a dollar in Euros revealed visually that an exponential functional relationship between the two variables may be appropriate. Therefore, the following exponential functional relationship;

$Y = AB^x$ is fitted to the data where:

Y = Gold dollar price

X = Value of a dollar in Euros

If this nonlinear equation was linearized as follows:

$$\ln Y = \ln A + (\ln B)x$$

The correlation analysis between Ln Y and X is also performed. The regression analysis and correlation analysis were conducted on daily and weekly data.

RESULTS

Daily data

Table 1

<u>Variable</u>	<u>Estimated Coeff</u>	<u>SE</u>	<u>t</u>
Constant	7.98	0.037	215.59
Dollar Value	1.98	0.0479	-41.38

Dependant variable = Ln of gold price
 Standard error of regression estimate = 0.03606.
 Adjusted R² = 79.4%

Model:

$$Y = (\text{Antilog of } 7.98) (\text{Antilog of } -1.98)^x$$

$$= (2922) (0.138)^x$$

Weekly data:

Table 2

<u>Variable</u>	<u>Estimated Coeff</u>	<u>SE of Coeff</u>	<u>t</u>
Constant	7.99	0.083	96.19
Dollar Value	-1.99	0.1075	-18.51

Dependant variable = Ln of gold price
 Standard error of regression estimate = 0.03625
 Adjusted R² = 79.3%

Model:

$$Y = (\text{Antilog of } 7.99) (\text{Antilog of } -1.99)^x$$

$$= (2951) (0.137)^x$$

2007, and it is possible that the relationship between the two variables studied could change over time.

FUTURE RESEARCH

Despite the evidence of a high negative correlation and exponential functional relation between the movements of dollar value and gold price in 2006 and 2007, it is important to investigate further to establish whether the current results hold true during different time period segments. It is also necessary to identify the other significant factors that may have influence on the gold price.

CONCLUSIONS

- During 2006 and 2007, there is about 90% negative correlation between dollar gold price and the dollar value in Euros.
- Adjusted R² is about 79% which means that about 21% of fluctuations in gold price are still unexplained. It shows that there are other significant factors which have influence on gold.
- The regression models fitted for the daily data, and for the data at the end of each week are same.
- The regression and correlation results are valid only for the years 2006 and

STOCK PERFORMANCE OF KIPLINGER'S RECOMMENDATIONS (2002-2003)

Thomas Pencek
University of South Florida—Sarasota/Manatee
8350 N. Tamiami Trail, SMC C263
Sarasota, FL 34243
941-359-4631
pencek@sar.usf.edu

ABSTRACT

This study investigates the financial performance of the Kiplingers' magazine. The magazine makes monthly stock picks for their readers. These picks are examined to determine if these picks would have outperformed a market portfolio. When the overall period (2002 through 2003) is examined, six-month returns yielded no abnormal returns. However, one-year returns of 7 percent were identified and they were significant. For New York Stock Exchange (NYSE) stocks, significance abnormal returns were observed. No significance from the market benchmark was observed for non-NYSE stocks.

INTRODUCTION

Kiplinger's is a monthly money management magazine that makes regular recommendations of stocks to purchase. *Kiplinger's* has almost 2.5 million subscribers. However, the number of subscribers is less than *Money* and *Smartmoney*. It has three main sections. Section one is the investing section. This is the section where money managers in *Kiplinger's* suggest stocks and mutual funds to invest in the future.

Barron's, *Business Week*, and the *Wall Street Journal* regularly publish columns containing expert opinions and rumors about firms. The content of these columns usually indicate positive or negative events or news that can have an impact on the share price of a company. For example, the *Wall Street Journal* publishes "Heard on the Street," *Business Week* has a section called "Inside Wall Street," and *Barron's* does it in "Up and Down Wall Street," in "Investment News and Views," in their Annual Roundtable, and in its featured articles. In contrast to the above magazines, *Kiplinger's* is published monthly. The price behavior and subsequent returns of magazines' recommendations have attracted the attention of researchers for two reasons: First, to see if the price behavior of these firms is affected because of the write-up and is of some potential value to investors; Second, to detect if there is any pattern to price behavior around the publication date that may be contradictory to market efficiency.

Market anomalies like insider trading, post-earnings announcement drift, small firm effect, January effect, day of the week effect, and price earnings effect have been well-documented. The purpose of this paper is to analyze the recommendations of those

in *Kiplinger's* magazine and observe the subsequent returns for these stocks. Furthermore, it takes a more long-term perspective of six months and one year.

This paper focuses on the small investor. *Kiplinger's* could be a source that a small investor could rely on. Although many options are available to this investor, a large number depend on the advice provided by *Kiplinger's*. The advice offered by *Kiplinger's* and other investment magazines is a reason to subscribe.

The Efficient Markets Hypothesis (EMH) suggests that all relevant, publicly available information is incorporated in the stock price of a company. An implication of the EMH is that excess returns cannot be earned. An anomaly of the EMH is the Value Line studies that found excess returns for certain ratings by Value Line. An examination such as this would expand the research for the EMH. Most of the recommendations are based on experts' opinions such as mutual fund managers.

BACKGROUND

Various authors have examined recommendations made by certain magazines. Most research centered on *Barron's*, *Business Week* and *the Wall Street Journal*. One study, by Brody and Rees (1995), has looked at recommendations by *Money* and *Changing Times* (now *Kiplinger's*). In that study, they observed 239 recommendations for 1990 by both magazines. They found that excess returns were not observed for those stocks recommended by the magazines after adjusting for the market. Some previous research follows.

Over the last 25 years, investors have relied on a stock rating system published by Value Line. Three studies, Copeland and Meyers (1982), Hall and Tsay (1988), and Pawlukiewicz and Preece (1991) found anomalies existed in the sense that above market returns are observed for stocks recommended by Value Line. These studies have found that stocks classified as a "one" in the timing rating outperform the market for the period that they have this rating.

Hemand, Desai, and Jain (1995) studied the effects of recommendations made by the participants in *Barron's* Annual Roundtable using 1968-1991 data. The Annual Roundtable is composed of invited financial analysts, fund managers, and others. A sample 1599 buy recommendations and 152 sell recommendations are identified. The authors analyzed the data for any abnormal returns associated with the recommendations over a short period of time, up to 10 days. For the positive recommendations, they found positive abnormal returns for the day of publication and that it was likely to be incorporated into the opening price. For the negative recommendations, they found significant abnormal returns on the day of distribution.

Trahan and Bolster (1995) examined the recommendations made in a variety of columns for *Barron's* for the year 1988. Their data consisted of 144 buy recommendations. They found significant positive returns of 2.1 percent on the day of the magazine distribution. They found positive, significant cumulative average returns from day 0 through day 16. The authors determined that articles written by Alan Abelson

had a more significant impact than others. Finally, they found that abnormal returns were more pronounced for smaller firms.

Sant and Zaman (1996) studied recommendations made in the *Business Week* column “Inside Wall Street” for the period 1976 through 1988. They recorded 456 cases of recommendations and categorized 40 negative, 88 inconclusive and 328 firms with positive recommendations. They found significant positive returns for three days surrounding the publication date for those stocks that had positive recommendations. For those stocks that received negative recommendations, no significance was identified. They also found significant returns before the publication date up to three months.

METHODOLOGY

The time period used for this study for stocks recommended by *Kiplinger’s* from January 2002 through December 2003. This period represents both a “bear” market period and a “bull” market period. A total of 247 stocks are recommended during this period: 170 in 2002 and 77 in 2003. Of this group, 179 stocks recommended trade on the New York Stock Exchange and 68 stocks are not traded on the New York Stock Exchange. These stocks are purchased based solely on the recommendations of *Kiplinger’s*.

Price information beginning with the month of publication is obtained and the prices six-month and one-year later are observed. The information is adjusted for any dividends or stock splits that occurred during that time. For example, if a stock is recommended in the January issue, prices are observed for that month and then the price six months later are observed. In addition, the price one-year later is observed. Any February recommendations are observed in a similar manner. Returns (R_t) are computed using annualized returns calculations. These returns are defined as R_t .

The returns are adjusted for the market by subtracting the market returns for the corresponding periods. These markets returns are computed by using the beta coefficient for the stock when the stocks are recommended. These beta coefficients are used to compute the market returns (R_m). Higher betas will cause the market returns to be higher. This is done to see if the recommendations “Beat the Market.” The abnormal returns (AR) are computed using the following formula:

$$AR = R_t - R_m$$

where R_m is the market return. As with other studies, the S&P 500 index is used as the barometer for the market.

These abnormal returns are summarized in terms of an average market adjusted returns. Furthermore, standard deviations are computed. These averages are compared to zero to observe any significant market returns. If positive and significant, it implies consistent above market returns. It would also imply to invest in these stocks when recommended and the investor should outperform the market. If not significant, it would imply that the investor would not have beaten the market if they invested in the stocks. If

this is the case, then the EMH would hold. The overall period is analyzed as such. In addition, each year is analyzed to determine if any significance is observed. For example, does a “bull” market have different stock price reactions than a “bear” market? 2002 would be considered to be a “bear” market and 2003 is considered to be a “bull” market.

Finally, the same analysis is applied to the stocks that trade on the New York Stock Exchange only. The overall period and each year are examined. Next, the same procedure is performed for those stocks that do not trade on the NYSE exchange. This analysis is performed to observe any differences in the price behavior of the different markets. For example, most NON-NYSE stocks are not followed as closely as those traded on the NYSE. Therefore, it could be argued that some significant returns may be observed on those stocks that do not trade on the NYSE.

RESULTS

All Stocks

On average, the six-month return was 2.5% percent for all stocks recommended by *Kiplingers*'. However, no significance was observed. A 7.0 percent abnormal return is observed for one-year returns. In addition, a very high level of significance is observed implying a consistent, positive above market returns if you held these stocks for one year.

Six-month abnormal returns by year are examined next. For 2002, a 0.6 percent rate of return is observed. It is not found to be significant. For 2003, a 6.6 percent rate of return is observed. It is significant at the five percent level. The implication here is that you would have consistently “beat” the market if you invested in the recommendations for 2003.

One-year returns for each year are examined next. Stocks recommended in 2002 had a 5.6 percent return. Furthermore, it was found to be significant implying there was a consistent above market return. A 10.2 percent return was observed for stocks recommended in 2003. Furthermore, it was found to be significant.

New York Stock Exchange Stocks

Stocks recommended that trade on the New York Stock Exchange only are examined next. Six-month returns have a positive 3.4 percent return and it is significant. For the one-year returns, a 6.7 percent rate of return was computed. In addition, it is found to be highly significant. For this study, it appears that investing only in NYSE stocks recommended that above market returns are achieved.

Examination of the six-month returns by each year is considered. In 2002, a 2.0 percent abnormal average return is observed. However, it is not found to be significant. In 2003, a 6.7 return is observed and it is significant at the 5 percent level.

One-year returns are examined next. In 2002, a 5.1 percent return is observed for those stocks traded on the NYSE recommended by Kiplinger's. In addition, it is found to be significant. In 2003, a 10.4 percent rate of return is observed. It is found to be significant. In both years, significant above market returns are found.

NON-NYSE Stocks

NON-NYSE stocks are examined next. Six-month returns found that, for the entire period, a negative 0.1 percent return is observed. However, it is not found to be significant. One-year returns are computed to be a 7.7 percent and it is not found to be significant.

In 2002, a negative 3.3 percent six-month rate of return is observed. However, it is not significant. In 2003, a 6.3 percent abnormal return is computed. However, it not found to be significant.

One-year returns for NON-NYSE are examined by years. In 2002, a 6.8 percent return is computed. However, due to high volatility, it is not significant. In 2003, a 9.6 percent return is computed. In addition, it is not found to be significant.

SUMMARY AND CONCLUSIONS

This study examined the effectiveness of an investment strategy based on analysts' recommendations published in *Kiplinger's*. For the overall period, one-year returns were found to be significant. When broken down by years, one-year returns were significant for both years. When only New York Stock Exchange stocks were examined, six-month and one-year returns were positive and found to be significant. However, when subdivided by year, 2002 was a year in which six-month returns were found to be not significant. NON-NYSE stocks had negative six-month abnormal returns for the entire period but they were not significant. When individual years were studied, none are found to be significant.

The evidence suggests mixed results. For all the stocks, recommendations made by *Kiplinger's* do not provide a reliable estimate of the future performance of a company's stock price. This would imply the EMH holds. However, the evidence for New York Stock Exchange stocks showed significant abnormal returns, particularly for 2003. This would suggest market inefficiencies. Given that more analysts follow the New York Stock Exchange more closely, this result is perplexing. These results suggest more recommendations should be followed to find similar results. The results also suggest that you should avoid any NON-NYSE stocks recommended by *Kiplinger's*.

BIBLIOGRAPHY

1. H. Baghestani and D. Nelson, "How Accurate Are Professional Economic Forecasts?" *Journal of Applied Business Research*, (Summer 1994), pp. 1-5.
2. R.G. Brody and L. Rees, "The Performance of Popular Investment Magazine

- Stock Analysts,” *Journal of Applied Business Research*, (Winter 1995/1996), pp. 42-48.
3. S.J. Brown and J.B. Weaver, “Using Daily Stock Returns: The Case of Event Studies,” *Journal of Financial Economics*, (March 1985), pp. 3-31.
 4. T. Copeland and D. Mayers, “The Value Line Enigma (1965-1978): A Case Study of Performance Evaluation Issues,” *Journal of Financial Economics*, (November 1982), pp. 289-321.
 5. H. Desai and P.C. Jain, “An Analysis of the Recommendations of the ‘Superstar’ Money Managers at Barron’s Annual Roundtable,” *Journal of Finance*, Vol. L. No. 4. (September 1995), pp. 1257-1273.
 6. T.W. Hall and J.J. Tsay, “An Evaluation of the Performance of Portfolios Selected From Value Line Rank One Stocks: 1976-1982,” *Journal of Financial Research*, (Fall 1988), pp. 227-240.
 7. Y.S. Huang and R.A. Walking, “Target Abnormal Returns Associated With Acquisition Announcements: Payments, Acquisition Form, and Managerial Resistance.” *Journal of Financial Economics*, (December 1987), pp. 329-349.
 8. I. Mathur and A. Waheed, “Stock Price Reactions to Securities Recommended in Business Week’s ‘Inside Wall Street,’ *The Financial Review*, Vol. 30, No. 3, (August 1995) pp. 583-604.
 9. O. Palmon, H.L. Sun, and A. Tang, “The Impact of Publication of Analyst’s Recommendations on Returns and Trading Volume,” *The Financial Review*, Vol 29, No. 3, (August 1994), pp. 395-417.
 10. J. Pawlukiewicz and D. C. Preece, “The Impact of Value Line Special Situations Recommendations on Stock Prices: Evidence From the Over-the-Counter Market,” *The Financial Review*, (November 1991), pp. 547-68.
 11. R. Sant and M. Zaman, “Market Reaction to Business Week ‘Inside Wall Street’ Column: A Self-Fulfilling Prophecy,” *Journal of Banking and Finance*, Vol. 20, (1996), pp. 617-643.
 12. E. Trahan and P. Bolster, “The Impact of Barron’s Recommendations of Stock Prices,” *Quarterly Journal of Business and Economics and Finance*, Vol. 34, No. 4 (Autumn 1995), pp. 3-15.

U.S. BOND RATE-INFLATION DYNAMICS: 1954-2002

Krishna M. Kasibhatla, North Carolina A&T State University, sudhakrishna@bellsouth.net
Luis Eduardo Rivera Solis, Dowling College, Long Island, riveral@dowling.edu
John Malindretos, Yeshiva University, Jnmalindre@aol.com

ABSTRACT

This study, employing cointegration and error-correction methodology, examines the long-run relationship and short-run dynamics between the inflation rate and the long-bond nominal yield in the U.S. during June 1954 to August 2002. Empirical results obtained strongly support a long-run equilibrium relationship between the two rates. This finding is in support of the original 'augmented Fisher hypothesis'. The estimated results of the error-correction model indicate Granger causation running from consumer price inflation to the long-bond yield, and also, deviations of the series from their steady paths, are corrected by a gradual decline of the inflation rate in response to the Federal Reserve's policy actions.

INTRODUCTION

The objective of this study is to gain an insight into the dynamic relationship between yields on 10-year Treasury bond, the current benchmark of long-term interest rate, and the long-term expected inflation rate in the U.S. It is widely believed that the long-bond rate contains information about expected future inflation rate. The dynamic relationship between the 10-year bond rate and expected inflation rate as well as the long-run equilibrium relationship is investigated using the popular cointegration and equilibrium error correction methodology. The forces that produce a long-run equilibrium relationship between two or more variables imply mean reversion and existence of an error correction mechanism that characterizes the dynamic relationship between the variables.

Darby [1], Feldstein [5], and Tanzi [16], in their empirical studies have shown that nominal long-bond interest rates, in the absence of tax adjustment, must change by more than the change in the expected rates of inflation, after tax, for the real rate to be invariant. Darby's [1] study suggested that nominal rate, unadjusted for taxes, should change by 1.3 to 1.5 times the change in the expected rate of inflation, a sort of 'augmented Fisher effect'. Studies by McDonald and Murphy [11], Mishkin [13], Wallace and Warner [17], Phylaktis and Blake [15], and Evans and Lewis [4] supported the general view that changes in nominal long-bond interest rates reflect fluctuations in expected inflation, but not to the extent of the full one-to-one Fisher effect. Fisher's [6] own study did not find a one-to-one relationship between changes in inflation and the changes in the nominal interest rates. Fisher attributed the lack of the strict one-to-one relationship to some type of 'money illusion'. We will revisit the issue using vector error correction (VEC) methodology.

The remainder of this study is organized as follows. Section 1 provides an overview of the relationship between the two variables expressed as the original Fisher [6] equation, and the time series econometric model used to investigate the dynamic interactions between the two variables concerned. The data, sample period, and the estimated results are presented in Section 2. Section 3 contains a summary and conclusions of this study and a comparison of the results with some recent studies.

OVERVIEW OF THE FISHER HYPOTHESIS AND THE EMPIRICAL MODEL

The Fisher hypothesis is that the nominal bond interest rate (i) adjusts to changes in the expected rate of inflation (p_t^e). This also implies the efficiency of the financial asset markets. Optimizing behavior of agents in asset markets requires that the nominal yield on bonds and expectations of inflation move in the same direction. As a result, changes in expectations of future inflation will lead to changes in the expected rates of real return on assets. In order to keep the real return on financial and real assets constant, the nominal rate of interest must adjust to the expected rate of inflation. So, in the long-run the nominal interest rate cannot stray away from the expected rate of inflation. The intuition behind the Fisher hypothesis is the concept of cointegration. That is, the two time series, i_t and p_t^e , are integrated of the same order and are also cointegrated so that r_t is stationary.

According to Engle and Granger [3], if two time series are integrated of the same order and are found to be cointegrated, the appropriate procedure to model the series is the vector equilibrium error correction model (VECM). However, if r_t is not stationary, then changes in r_t will be a source of change in both the expected inflation rate and the nominal interest rate. Fisher's implicit assumption in (1.1) appears to be that i_t and p_t^e are integrated of the same order and also cointegrated while r_t is assumed to be stationary.

The original Fisher hypothesis stipulates a one-to-one relationship between expected rate of inflation and nominal interest rates, unadjusted for taxes. This hypothesis has been extensively tested in the literature with mixed empirical results. Evidence in some of the empirical studies, mentioned earlier, rejected the strict Fisher effect, but found less than complete adjustment of nominal interest rate to changes in inflation rates. This finding of less than full adjustment of nominal interest rates in some studies is largely attributed to the downward bias of the coefficient estimates caused by inadequate modeling or misspecification of the stochastic features of the data generating process, and to the sensitivity of the estimated results to the selection of the sample period and the country concerned (Widemann [18]). Widemann [18] questioned whether this relationship should be tested as a long-run equilibrium or cointegration relationship or within a threshold cointegration framework. According to Widemann [18], it is not clear whether inflation and nominal interest rates are integrated of the same order. However, when tested within the threshold cointegration, the Widemann's results perfectly supported the full Fisher hypothesis.

We will revisit this empirical question using a different but longer sample period, and by employing cointegration and vector error correction estimation procedure. Empirical results of this study will be compared with the earlier studies in the next section. Given the inconclusive nature of this relationship, and the mixed results reported in earlier studies, it is hoped that this study will make a good contribution toward resolving this issue.

THE EMPIRICAL MODEL

The long-run equilibrium relationship between changes in nominal interest rates and expected inflation is examined employing the Johansen and Juselius (JJ) [7],[8], maximum likelihood estimation procedure. If two or more time series share a common stochastic trend, then, the series are cointegrated. The presence of cointegration relation forms the basis of the VEC specification. Consider a vector autoregression (VAR) model of order p :

$$(1) \quad y_t = \mu + \sum_{i=1}^p A_i y_{t-i} + \varepsilon_t$$

where, y_t is a column vector of variables, μ is a vector of constants, and ε_t is a vector of innovations, assumed to be contemporaneously correlated but not auto correlated, and p is the number of lags of variables in the system.

If the variables in y_t are integrated of, say, order one, $I(1)$, and are also found to be cointegrated, that cointegration restriction has to be incorporated in the VAR in (2.1). The Granger Representation Theorem (Engle and Granger, 1987) states that variables, individually driven by permanent shocks are cointegrated, if and only if, there exists a VEC representation of the concerned time series data. A VAR model, with this restriction imposed, is referred to as VEC. Variables in the model enter the equation in their first or second differences, depending on the order of their integration, and the error correction terms are added to the model. So the VEC has cointegration relations built into the specification so that it restricts the long-run behavior of the endogenous variables to converge to their cointegrating relationship while allowing for short-run dynamics. Deviations from long-run equilibrium are corrected through a series of partial short-run adjustments.

The VEC representation of the VAR in (2.1), following \mathbb{J} is:

$$(2) \quad \Delta y_t = \mu + \sum_{i=1}^p \Gamma_i \Delta y_{t-i} + \Pi y_{t-1} + \xi_t$$

where, y_t is a $n \times 1$ vector of $I(1)$ variables. Γ_i is a $n \times n$ matrix of coefficients of the short-run dynamic effects, and, Π is a $n \times n$ matrix of coefficients of long-run effects, and ξ_t is a vector white noise process.

If the rank of Π in (2) is r , where $r = n-1$, then, Π can be decomposed into two $n \times r$ matrices, α and β , such that $\Pi = \alpha \beta'$. The matrix β is the cointegrating matrix of r cointegrating vectors, $\beta_1, \beta_2, \dots, \beta_r$. The β vectors represent estimates of the long-run cointegrating relationship among the variables in the system. The error correction terms, $\beta' y_{t-1}$, are the mean reverting weighted sums of cointegrating vectors. The matrix α is the matrix of error correction coefficients, the so called 'speed of adjustment' coefficients that measure the speed at which the variables adjust to their long-run equilibrium values.

If the rank of Π in (2.2) is found to be $r = n-1$, equation (2.2) can be expressed in the first differences of y_t , augmented by the error correction terms, $\alpha \beta' y_{t-1}$, as shown below:

$$(3) \quad \Delta y_t = \mu + \sum_{i=1}^p \Gamma_i \Delta y_{t-i} + \alpha \beta' y_{t-1} + \xi_t$$

The JJ technique provides maximum likelihood estimates of α and β' . The cointegrating relationship, r , is determined by the trace eigenvalue statistic and the maximum eigenvalue statistic of the stochastic matrix and the maximum likelihood estimates of the cointegrating vectors (β) in the equation (2.3).

The estimated model is a three variable VEC, and the variables are the 10-year Treasury long-bond rate, CPI inflation rate, and the effective federal funds rate.

DATA, MODEL ESTIMATION AND RESULTS MAJOR HEADING

Monthly data from 1954.07 to 2002.08 on the 10-year bond interest rate, the actual consumer inflation rate (CPI), and the effective monthly federal funds rate were obtained from the website of the Federal Reserve Bank of St. Louis (FRED).

Test results, presented in Table 1, indicate that the null hypothesis that each of the series in levels contained one unit root could not be rejected for any of the series. Then, unit root tests are performed on the first differences of series Δi_t , Δp_t^e and ΔFFR . The null hypothesis of a unit root could be rejected for each of the time series in their first differences at 1% level of significance.

Table 1: Augmented Dickey-Fuller (ADF) Unit Root Test:

Time Series	Lags	Intercept	Intercept & Trend	No intercept or Trend	Test Statistic	Akaike Information Criterion (AIC)
i_t (10-yr bond rate)	2		v		-1.6469	1.3435
Δi_t	2		v		-10.8794*	1.3473
p_t^e (inflation rate)	2	v			-2.7458	-0.0501
Δp_t^e	2	v			-16.0017*	-0.0217
FFR (Fed Funds Rate)	7	v			-1.9904	2.6443
ΔFFR	6			v	-6.4026*	2.6449

The results of cointegration tests are presented in Table 2. In both the cases the trace test (which tests the null hypothesis of n cointegrating relations against k cointegrating relations, where k is the number of endogenous variables, for $n = 0, 1, \dots, k$), indicates two cointegrating vectors at 5% level of significance. The maximum eigenvalue test (tests the null of r cointegrating relations against the alternative of $n+1$ cointegrating relations) results indicated two cointegrating equation at the 5% level of significance.

Table 2: Cointegration Test Results: i_t , p_t^e , and FFR. 1954:11 to 2004:08.

Assumptions: Intercept and deterministic trend in CE and no intercept in VAR. Lags interval (in first differences): 1 to 4, based on the Akaike Information Criterion.							
Hypothesized # of cointegrating equations	Eigenvalue	Trace Statistic	5% Critical Value	1% Critical Value	Max-Eigenvalue Statistic	5% Critical Value	1% Critical Value
None	0.1052	135.75	42.44	48.45	63.172	25.54	30.34
At most 1	0.1033	72.58	25.32	30.45	61.947	18.96	23.65
At most 2	0.0185	10.63	12.25	16.26	10.634	12.25	16.26
Trace Test and Maximum Eigenvalue Test indicate 2 cointegrating equations at 1% level.							
Critical values: Osterwald-Lenum(1992)							
Normalized cointegrating coefficients (std.err. in parentheses)							
i_t	p_t^e	FFR	C	Log Likelihood			
1.0	-1.287 (0.0901)	-0.866 (0.0070)	0.006 (0.0021)	2120.872			

The normalized cointegrating coefficients, in Table 2 for the long-bond rate and inflation indicate a more than a one-to-one relationship (1 to 1.287) between nominal interest rate and inflation rate (p_t^e). Our results are close to the results suggested by Darby [1] between 1.3 and 1.5 rather than to Mehra's 1996 study. However, our cointegration test results are consistent with Mehra's [12] findings that these three series are cointegrated, and there is a long-run equilibrium relationship among the three series.

Table 3: VEC Estimation Results. 1954.10-2004.08

Vector Error Correction Estimates			
t-statistics in []			
Cointegrating Eq:	CointEq.1	CointEq.2	
	$\Delta yr10$	Δp_t^e	
Error Correction:	$\Delta yr10$	Δp_t^e	ΔFFR
Coint.Eq.1	-0.817 [-7.648]*	0.0016 [1.602]	1.1933 [12.198]*
Coint.Eq.2	0.221 [2.164]	-0.005 [-0.569]	-0.169 [-1.563]
yr10 _{t-1}	0.169 [1.998]	-0.0004 [-0.054]	-0.2958 [-1.821]
yr10 _{t-2}	0.1428 [1.943]	0.002 [0.225]	-0.361 [-2.562]
yr10 _{t-3}	-0.111 [-0.194]	-0.003 [-0.466]	-0.362 [-2.492]
yr10 _{t-4}	-0.109 [-2.261]	[0.002]	-0.183 [-2.067]
Δp_{t-1}^e	2.673 [4.331]	-0.771 [-0.042]	-8.469 [-2.306]
Δp_{t-2}^e	0.812 [2.015]	-0.423 [-0.864]	-2.698 [-0.276]
Δp_{t-3}^e	-0.541 [-1.965]	-0.356 [-0.278]	-7.054 [-1.980]
Δp_{t-4}^e	1.231 [2.021]	-0.177 [-0.431]	-6.112 [-1.407]
ΔFFR_{t-1}	0.007 [1.928]	-0.081 [-0.094]	0.342 [1.876]
ΔFFR_{t-2}	0.045 [0.985]	-0.008 [-0.967]	1.23 [2.043]
ΔFFR_{t-3}	0.146 [1.654]	0.184 [0.761]	0.591 [0.853]
ΔFFR_{t-4}	1.29 [0.732]	0.651 [0.147]	-0.142 [-0.543]
C	-0.006 [-0.053]	0.002 [0.171]	-0.011 [-0.053]
R-squared	0.4886	0.4637	0.4105
Adj.R-squared	0.4437	0.4277	0.3957
S.E. equation	0.2609	0.0025	0.5004
F-statistic	27.4873	22.7397	27.7067
Log Likelihood		2205.344	
Log Likelihood (d.f. adj.)		2182.544	
Akaike Infor. Criterion		-7.4459	
Schwarz Criterion		-7.0430	

The key aspects of interest for this study are the coefficients of the normalized cointegration equations, the coefficients of the error correction terms and the lagged values of the variables presented above in Table 3.

SUMMARY AND CONCLUSIONS

The empirical results of our study strongly indicate that bond rate and consumer inflation rate are cointegrated. Secondly, bond rate is caused by consumer price inflation and consumer inflation is not caused by bond rate. Further, our empirical results support Darby's finding of 'augmented Fisher effect.' The implication of this finding is that interest rate, not adjusted for taxes, has to increase by nearly 1.28 when the inflation rate rises by one unit in order to keep the real rate constant. However, these findings are not in agreement with Mehra's [12] study.

REFERENCES

Referentes provided upon request.

GLOBALIZATION: A CONSUMER BEHAVIOR PERSPECTIVE

Alberto Rubio-Sanchez, Roger Williams University. One Old Ferry Road, Bristol, RI 02809.
E-mail: alrubio@rwu.edu. Phone: 401-254-5668

ABSTRACT

The background culture of an individual has always been considered a big influence in its shopping and consumption behavior. Due to recent globalization trends, the differences in culture have been diminishing, leading practitioners and researchers to wonder if the behaviors of individuals from different cultures are also becoming more homogeneous. The present study uses the World Values Survey 1981-1995 to show that cultural gaps are, in fact, closing. A larger study is proposed to investigate if globalization is truly taking place, and if so, how would it impact the consumption behavior of individuals around the world.

INTRODUCTION

In the consumer behavior literature, cultural and national values have been recognized as a powerful force shaping consumers' motivations, life-styles, and product choice [1] [6] [8] [11]. However, new evidence shows the differences in values and cultural dimensions may be a thing of the past [10]. Some authors even posit the argument that cultural differences have become overrated and insignificant [7]. If these assertions are true, it would mean, to the very least, that we need to change the way we market our products and services in other cultures.

Eight years into the twenty first century we can observe that industrialization, proposed by Carl Marx as the main source of cultural change, is not the only force closing the gap between cultural values of different nations. Globalization activities, especially by American companies, have reached every square inch of the planet, bringing American cultural values to the world [5]. These American values and tastes are usually transferred to other nations through media, political influence, and various foreign activities of their citizens and companies [9]. In our time this has been sometimes called the "McDonalozation" of the world [5].

The topic of globalization, and not only "Americanization," has received considerable attention in recent years. America is not the only national power to export their values to the world. Expansion of world travel, declining trade barriers, advances in communication technology, the internet, and the emergence of global media are all factors contributing to the acceleration of a completely globalized market [4].

To explore if the globalization of values is a reality, a study with the World Values Survey (1981-1990-1995) was used. This data series is designed to enable a cross-national comparison of values and norms on a wide variety of topics and to monitor changes in values and attitudes across the globe. A very broad range of topics is covered in the series.

Of special interest for the present study were those topics relating to the cultural dimensions proposed by Hofstede [2]: Individuality, Uncertainty Avoidance, Masculinity-Femininity, and

Power Distance. Due to the secondary nature of the data only three variables seemed to possess enough face validity to be treated as elements of the four original cultural dimensions. Also, only two countries were included in this preliminary analysis: the United States and Mexico. These countries were chosen because, traditionally, they have been on opposite ends of the cultural dimension spectrums being studied here, thus providing the researcher with enough material to test the theory.

Some of the preliminary results of the study are encouraging. As an example, three of the compared items are presented. These items pertain to the cultural dimensions previously mentioned.

The first dimension is individualism. One of the characteristics of individualist countries is their partiality for independence from the group. In one of the questions, the WVS presents respondents with a list of 10 characteristics and lets them choose five that they feel are important to teach their offspring. One of those characteristics is independence.

Univariate analysis reveals that significant changes have occurred in the way respondents from the two countries looked at independence in 1981 and 1995. In 1981, 31% of American respondents thought it was important to teach children to be independent, compared to only 15% of Mexicans ($\chi^2 = 109.8$, $df = 1$, $p < .000$). In 1995, however, 45% of Americans thought independence was important, and 40% of Mexican respondents felt the same way ($\chi^2 = 9.1$ $df = 1$, $p = .003$).

We can clearly observe that the gap between the United States and Mexico in the way they look at independence, changed significantly in only 14 years. This is consistent with the proposed theory showing a preference for a more independent, perhaps individualistic, society.

The second dimension studied was power distance. The WVS asks how important is to teach obedience to children. Obedience and respect are two of the main pillars of power distance. Once again, the data seems to support the theory. In 1981, 26% of American respondents thought it was important to teach children to be obedient, compared to 44% of Mexicans ($\chi^2 = 105.9$, $df = 1$, $p < .001$). In 1995, 38% of Americans thought obedience was important, and 49% of Mexican respondents felt the same way ($\chi^2 = 46.1$ $df = 1$, $p < .001$). Even though the difference in opinions of respondents from the two countries is still significant, the narrowing of the gap is unmistakably apparent.

Lastly, we looked at the third dimension: uncertainty avoidance. WVS asks:

Some people feel they have completely free choice and control over their lives, while other people feel that what they do has no real effect on what happens to them. Please use this scale where 1 means "none at all" and 10 means "a great deal" to indicate how much freedom of choice and control you feel you have over the way your life turns out.

This question taps into the essence of uncertainty avoidance [3].

After the analyses were conducted we can observe that in 1981 the mean value for the American respondents was 7.44 ($\sigma = 2.1$), while the mean for the Mexican sample was 7.58 ($\sigma = 1.9$). In 1995, the mean for the Mexican sample remained almost unchanged ($M = 7.56$, $\sigma = 2.3$), but the American sample mean varied to 7.54 ($\sigma = 2.0$).

Contrary to the what was observed in the other two dimensions, American respondents seemed to significantly change their perception of how much control one has over one's life ($F = 3.3$, $p = .06$), while the Mexican respondents' perception remained unchanged ($F = .01$, $p = .89$). None the less, the results continue to show a narrowing of the gap between the values and perceptions of both countries.

The conclusions reached with this preliminary study provide the researcher with enough confidence to proceed with a broader study of four countries. Also, the use of primary data using measures that more accurately reflect values influencing shopping and consumption behavior seems in order. The analyses proposed for this study include structural equation modeling and latent growth curves.

REFERENCES

- [1] Carman, J. M. "Values and consumption patterns: a closed loop." *Advances in Consumer Research*, 1978, 5 (4), 403-407.
- [2] Hofstede, G. H. *Culture's consequences*, Thousand Oaks, CA: Sape Publications Inc, 1980.
- [3] Hofstede, G. H. *Culture's consequences*, Thousand Oaks, CA: Sape Publications Inc, 2001.
- [4] Hofstede T. Steenkamp J.B.E.M. Wedel, M. "International market segmentation based on consumer-product relations." *Journal of Marketing Research*, 1999, 36 (1), 1-17.
- [5] Inglehart, R. Baker W. E. "Modernization's challenge to traditional values: Who's afraid of Ronald McDonald." *The Futurist*, 2001, 35 (2), 16-24.
- [6] Kluchohn, C. "Value and value orientations in the theory of action." in *Towards a General Theory of Action*, Eds. Talcott Parsons and Edward.A. Shils, Cambridge, MA: Harvard University Press, 1969, 388-433.
- [7] Markoczy, L. "National culture and strategic change in belief formation." *Journal of International Business Studies*, 2000, 31 (3), 427-442.
- [8] McCracken, G. D. "Culture and consumption: A theoretical account of the structure and movement of the cultural meaning of consumer goods." *Journal of Consumer Research*, 1986, 13, 71-84.
- [9] Porter, M. E. "Clusters and the new economics of competition." *Harvard Business Review*, 1998, 76 (6), 77-90.
- [10] Pressey, A. D. Selassie, H. G. "Are cultural differences overrated?" *Journal of Consumer Behavior*, 2003, 2 (4), 354-368.
- [11] Yau, O.H.M. "Chinese culture values: Their dimensions and marketing implications." *European Journal of Marketing*, 1988, 22 (5), 44-57.

Aligning International Business, Human Resources & Information System Strategies

Arun Madapusi
University of North Texas
Department of Management, College of Business Administration
P.O. Box 305429, Denton, Texas 76203.
Phone: (940) 565 3338; E-mail: madapusi@unt.edu

Abstract

Past studies indicate that firms need to align their international business and human resource strategies to achieve superior performance. As firms rollout information systems (IS) across their global operations it is also imperative that firms align their IS with their business and human resource strategies. This research study examines the international business-human resource-IS strategy alignment issue and develops a framework to guide managers in aligning IS with their business and human resources strategies.

Keywords: Business Strategy, Human Resources Strategy, Information Systems, Alignment.

Introduction

Globalization of trade and economics has forced firms to seek internationally experienced managers to run their worldwide operations. Researchers indicate that a key ingredient for firms in successfully implementing international business strategies is the adoption of compatible international human resource strategies [3] [7]. As firms seek to match their human resources and business strategies, a crucial issue that has emerged over the past decade is the impact of global information systems (IS) on business strategy-human resources alignment. This research study examines the above issue and develops a framework to guide managers in aligning IS with their business and human resources strategies.

Aligning International Business and Human Resource Strategies

Studies indicate that organizational forces influence a firm's international business strategy [2]. A firm's value-chain activities are impacted by two environmental forces – national differentiation and global integration. Firms follow four sequential strategies in the path to the internationalization of their operations – from multinational to international, then to regional, and finally to transnational [2]. Researchers indicate that firms typically follow four international human resources strategies – ethnocentric, polycentric, regiocentric, and global [1] [4]. Studies that examined the alignment between business and human resource strategies indicate that ideal

fit combinations lead to increased performance benefits [3] [7]. Multinational firms typically prefer polycentric strategies; international firms use ethnocentric strategies, regional firms adopt regiocentric strategies, and transnational firms opt for global strategies [3] [7].

Aligning Information System, International Business and Human Resource Strategies

Studies indicate that firms can configure IS to support their business strategies [5] [6]. The choices of configurations include software configurations, technology platforms, and implementation approaches. The software configuration is mainly undertaken at four levels [5] – enterprise, system, business process, customization. At the enterprise level firms can configure their IS according to their business strategies. At the system level core business activities are implemented as modules and rows of sub-modules within each of these activities. The business process level focuses on the customization of user profiles, parameters, and business processes at various organizational levels. The customization level involves custom-designed modifications of the IS. The technology platform involves the use of centralized, decentralized, or hybrid configurations [9]. There are also two broad implementation approaches that firms typically use to deploy their IS – global and phased rollouts [8].

The configurational choices made at each of the three levels – software configuration, technology platform, implementation approach – determines the alignment of the IS with the firm's international business strategy-human resources strategy fit, and hence firm performance. A multinational-polycentric strategy requires relatively independent software configurations, local databases, and phased rollouts. An international-ethnocentric strategy, on the other hand, needs centralized software configurations, centralized databases, and global rollouts. The software configurations, databases, and rollouts for a regional-regiocentric strategy are similar to a multinational-polycentric strategy but on a larger scale. The transnational-global strategy entails distributed software configurations, hybrid databases, and hybrid rollouts. At each of the three IS configurational levels, firms can drill down to detailed levels and customize their IS to unique and firm-specific human resource needs. This detailed customization; however, should be within the broad gambit of the overall business strategy-human resource strategy-IS configurational fit. Firms that successfully manage the above alignment process can achieve superior performance in the global marketplace.

References

- [1] Adler, N.J., Fariborz, G. "International Strategy from the perspective of people and culture: The North American context." *Research in Global Business Management*, 1990, 1, 179-205.
- [2] Bartlett, C.A., Ghoshal, S. *Managing Across Borders: The Transnational Solution*, Boston: Harvard Business School Press, 1998.

- [3] Cullen, J.B., Parboteeah, K.P. *Multinational Management: A Strategic Approach*. 3rd. Ed. Ohio: South-Western, 2005.
- [4] Heenan, D.A., Perlmutter, H.V. *Multinational Organization Development*, Reading, Mass: Addison-Wesley, 1979.
- [5] Koch, C. "Enterprise Resource Planning: Information Technology as a Steamroller for Management Politics." *Journal of Organizational Change Management*, 2001, 14(1), 64-78.
- [6] Markus, M.L., Tanis, C., Van Fenema P.C. "Multisite ERP Implementations." *Communications of the ACM Journal*, 2000, 43(4), 42-46.
- [7] McFarlin, D.B., Sweeney, P.D. *International Management: Strategic Opportunities and Cultural Challenges*. Boston: Houghton Mifflin Company, 2003.
- [8] Parr, A., Shanks, G. "A Taxonomy of ERP Implementation Approaches." *Proceedings of the 33rd. Hawaii International conference on System Sciences*, 2000, 1-10.
- [9] Zrimsek, B., Prior, D. "Comparing the TCO of Centralized vs. Decentralized ERP." *Gartner Research*, Gartner Inc., 2003

INFLUENCE OF INTERNET MARKET ORIENTATION ON SHOPPING WEBSITES SUCCESS : AN EMPIRICAL STUDY OF ONLINE SHOP BUSINESS IN TAIWAN

Wei-shang Fan, Nan Hua University, (886)911-111243, wsfan@mail.nhu.edu.tw
Ruei-Shiuan Chang, Nan Hua University, (886)928-789797, rschang@cht.com.tw
Lily Hwang, Nan Hua University, (886)921-519828, cherry8645@yahoo.com.tw

ABSTRACT

This research aims to probe into the influence of Internet market orientation on successful cases of online shop. The research objects were the top three shopping websites (e-bay, Yahoo!, and PChome). Factor analysis and Structure Equation Model (SEM) were used to verify the overall model. The exploration is based on questionnaire survey, and the research applied contingency approach to analysis the results. The combination of promotional strategies becomes an important tool of Internet marketing, thus, requires proper and cautious planning. This research explores the influence of contingency and structure on decision-making results during the decision-making process. The empirical results showed that Internet market orientation can positively influence Internet marketing strategies, Internet customization and website design quality; Internet marketing strategies and website design quality can positively influence successful cases; Internet customization can positively influence website design quality; website design quality can positively influence Internet marketing strategies.

Keywords: Internet market orientation, Internet marketing strategies, Internet customization, Website design quality, Successful cases

INTRODUCTION

With vigorous development of the Internet and growth of communication technology, e-Commerce not only has changed the traditional business transaction model, but also led to enormous business opportunities. Companies can conveniently collect information on the market and customers through various channels, and establish customer related database, thereby, effectively deliver and disseminate information to internal and external members of the companies, so that company can respond customers' demands through promotions (Min, Song & Keebler, 2002).

As the Internet provides more and more comprehensive information, website design has become more and more complicated. Some common problems associated with online shopping websites include difficulty to find detail specification of the product even though it is listed on the webpage, lack of information classification and repetitive information. Because of these problems, users usually could not find the information they want, and waste too much time. Thus, online shopping websites should consider how to provide users with individualized information and service, and allow them to find the information they need by improving the efficiency of search.

LITERATURE REVIEW AND HYPOTHESES

This research aimed to construct a success model for shopping websites. The application of contingency theory framework followed the contingency model of contingency->structure->result

proposed by Fiedler (1967). He emphasized that in order to respond to different contingencies, companies must adjust their structures and strategies to obtain optimized results. Based on the above concepts of social system, we confirmed the cause and effect of website success, and constructed the contingency model, this research first considered Internet market orientation because market orientation created more consumer value and improved corporate performance (Sanzo, Santos & Vazquez, 2003; Narver and Slater, 1995; Jaworski & Kohli, 1993).

With regard to result construct, that only the websites providing customized service and positive design quality could increase customer satisfaction and repurchase intention to achieve website success. This research mainly assessed customer satisfaction and trust toward shopping websites after consumption, as well as their repurchase intention, for the basis of evaluating website success. Based on above, we constructed this research framework, as shown in Figure 1.

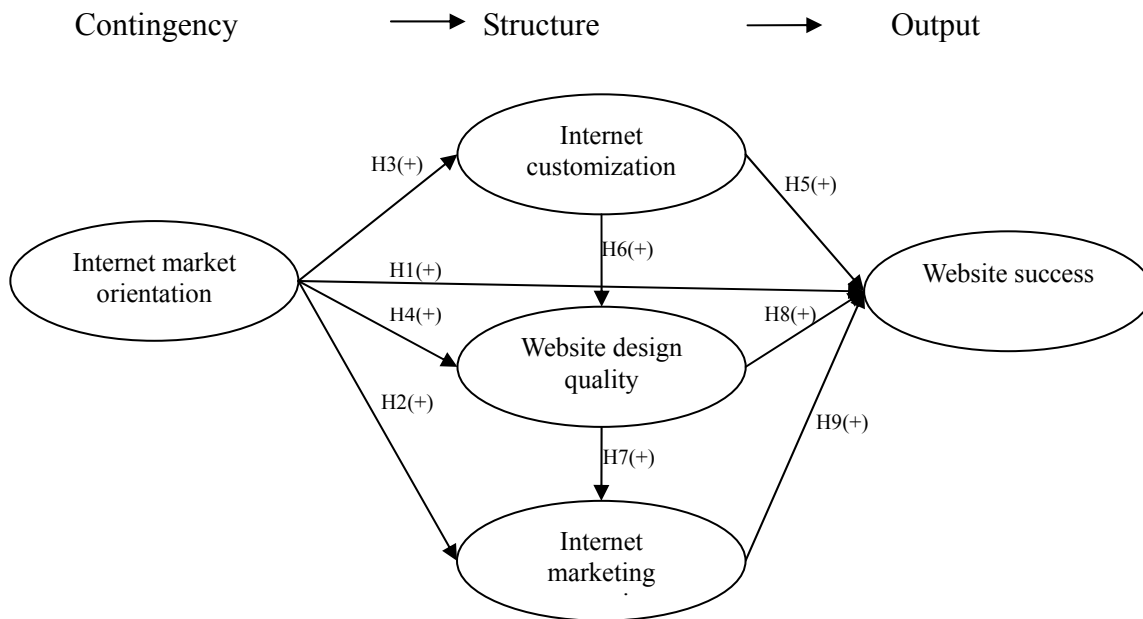


Figure 1 Research framework

Three core factors of Internet market orientation include producing information needed by customers, processing and diffusion, collecting customers' information, establishing customer database, and reacting to customer demand. The main concept emphasizes that the companies should understand customers' true needs, as well as advantages and disadvantages of their rivals, then create values for customers through coordination of corporate resources and marketing strategies (Min, Song & Keebler, 2002). Internet challenges the operation of traditional market, makes prediction based on past operational records, and influence corporate and industry performance and marketing strategies completely. Kotler (2000) segmented consumer groups according to their purchasing characteristics or reactions, and designed marketing strategies combinations based on special needs and expectation of various consumer groups toward online shopping. Innovation and integration of new service technology play critical roles in market orientation and organizational performance. The integration of Internet technology and marketing could increase companies' competitive advantages of marketing-related competence.

H-1: There is significant and positive correlation between Internet market orientation and website success.

H-2: There is significant and positive correlation between Internet market orientation and Internet marketing strategies.

H-3: There is significant and positive correlation between Internet market orientation and Internet customization.

H-4: There is significant and positive correlation between Internet market orientation and website design quality.

Because of the rapid development of Internet functions and aggressive development of customized services by website companies, online shopping activities have gradually focused on providing online shoppers individualized and convenient services to meet their needs. Because of the needs to search for shopping information through Internet, shopping websites have actively developed search function to maintain customer loyalty and establish competitive advantages (Srini et al., 2002). Experts suggested that Internet users value the convenience of webpage interface and individualized website design. Besides, functions such as customized search interface, user-friendliness, and substitute for physical contact could satisfy users' needs for usability and enjoyment, hence, further strengthen their purchase intention. Burke (1997) suggested that strengthening customized services for shopping websites could be conducive to online users in searching product information. Szymanski and Hise (2000) found that the principal factors of American consumers' satisfaction toward online shopping are convenience, website design, and financial security. According to above literature, this research proposed the following hypotheses:

H-5: There is significant and positive correlation between Internet customization and website success.

H-6: There is significant and positive correlation between Internet customization and website design quality.

Lee & Kozar (2006) explored the factors affecting website quality, and argued that when companies select and design more flexible websites, they could achieve corporate financial and website performances. DeLone and McLean (2003) proposed the IS success model, which divided website quality into information quality, system quality, service quality, and vendor-specific quality. Their research findings showed that higher the website quality, better the corporate performance. In other words, website quality directly affects corporate performance. Liu and Arnett (2000) indicated that for website designers, websites must be able to cultivate customers' interests, enhance their participation intention, and facilitate their excitement when using the websites. In addition, websites should have appealing characteristics to attract customers and allow them to enjoy browsing the websites. Szymanski & Hise (2000) argued that website design has significant and positive influence on Internet customer satisfaction. According to above literatures, this research developed the following hypotheses:

H-7: There is significant and positive correlation between website design quality and Internet marketing strategies.

H-8: There is significant and positive correlation between website design quality and website success.

Online shoppers' purchase behavior is affected by marketing strategies of the website companies. When Internet users obtain product information through Internet, companies can attract the consumers by more competitive prices. Therefore, Internet consumers with higher price sensitivity would be attracted by low-price strategies and their purchase intention would also increase (Jiang Mei-Hsiang, Yeh Kuei-chen, 2000). The rise of online shopping consumption is due to reduced channel cost and lower price, so that consumers build trust toward the shopping websites and develop purchase intention (Korgaonkar, 1984). Internet marketing is how a company focuses on target customers and provides service products or product information through Internet, then acquires analytical data and builds customer relation to accomplish mass sales. (Kalakota & Whinston, 1996). Therefore, this research developed the following hypothesis:

H-9: There is significant and positive correlation between Internet marketing strategies and website success.

RESEARCH DESIGN

Variable definitions and assessment are based on empirical results of related researches. Besides considering reliability and validity of the questions of original scales, we also generalized industry characteristics, environment and meanings of questions.

Questionnaire design of this research was based on literature review and questionnaire design of the related studies, and was revised and translated based on the requirements of this research. The draft of the questionnaire was reviewed by experts, and revised according to their opinions before entered into pre-test. The questionnaire was loaded onto <http://www.my3q.com/>, and invitations letters were sent to blog users to fill in the questionnaire on the website. The questionnaire was distributed from early January 2007 to early February 2007. A total of 1000 questionnaires were distributed, and 269 were returned. After eliminating 49 questionnaires with incomplete and repetitive answers, there were 220 valid questionnaires; the valid return rate was 22.0 %.

RESULT OF ANALYSIS

This research examined the model and the result showed that the indices of overall sample model all reached the assessment criterion, as shown in Table 1. All the fit assessment indices were close to the criteria, except for GFI and AGFI of less than once actual transaction per month. Therefore, generally speaking, among overall samples of this research, the sample fit of average actual transactions of over once, once, and less than once per month reached the goodness of fit.

Table 1 Analytical result of model fit

Fix assessment indices	Overall sample model (n=220)	Actual shopping for once and at least once per month in average (n=99)	Actual shopping below once per month in average (n=121)	Judgment principle
χ^2 value (DF)	282.4(138)	157.67(136)	235.65(138)	The smaller the better
χ^2/DF	2.122	1.159	1.708	<3
P value	0.000	0.0985	0.000	Not significant
GFI	0.89	0.87	0.83	>0.9
AGFI	0.84	0.82	0.78	>0.9
RMR	0.032	0.034	0.044	<0.05
RMSEA	0.069	0.040	0.076	<0.08
NFI	0.91	0.91	0.86	>0.9
CFI	0.95	0.99	0.94	>0.9
PNFI	0.79	0.72	0.61	>0.5
PGFI	0.64	0.62	0.69	>0.5

Note: *P<0.05 **P<0.01 ***P<0.001

CONCLUSIONS

This research validated the overall model of Internet market orientation, marketing strategies, customization, Internet marketing strategies and website success, as shown in Figure 2. The results showed that Internet market orientation could directly and positively influence Internet marketing strategies, customization and website design quality; Internet marketing strategies and website design quality could directly and positively influence website success; customization could directly and positively influence website design quality; website design quality could directly and positively influence Internet marketing strategies; however, only customization cannot directly influence website success. According to research observation, at present, the shopping websites in Taiwan have not developed complete customized service model from customers' product ordering and price estimation. Therefore, product customization process cannot be established as in the technology industry, such as Dell. Finally, fit test of linear structural relationship model showed that overall theoretical model fit of this research met the test criterion, indicating that the theoretical model of

this research was supported.

Management Implications

Building shopping websites has become a popular marketing practice. Internet and technology information play significantly important roles in the global economy. With the development of technology, e-Commerce becomes a mainstream economic body. Internet channel also have changed consumers' living and consumption habits so that companies are facing inevitable reform and challenge. From the view of business practice, the companies must value the planning of Internet market orientation and establishment of Internet marketing strategies to attract consumers to their websites and make purchases. Another key to successful website operation is that the companies should provide complete shopping information, products, and services needed by individual shoppers to reduce their search time and allow them to be satisfied with the services.

More importantly the companies should value the overall design of websites and provide complete product information to and concern about the security of website transaction to avoid the exposure of customers' confidential data. In addition internal marketing related departments of the companies should communicate with other departments with focus on meeting customers' needs and forming strategies accordingly.

Internet is a changeable and future development, and changes will lead to new affecting factors at any time. Future researches can include new variables in their research framework and compare different industries in this model.

REFERENCE

1. Burke, Raymond R., Virtual Shopping: Breakthrough in Marketing Research, *Harvard Business Review*, 1997,74 (March-April), pp.120-131.
2. DeLone W.H., McLean E.R., The DeLone and McLean model of information systems success: a ten-year update, *Journal of Management Information Systems* ,2003, 19, pp. 9-30.
3. Fiedler, F. E., *Contingency theory*, New York: McGraw-Hill, 1967.
4. Kalakota, R. and A.B. Whinston, *Frontiers of Electronic Commerce*, Addison- Wesley Company,1996.
5. Kohli, Ajay K., Jaworski, Bernard J, Kumar, Ajith, MARKOR: A measure of market orientation, *Journal of Marketing Research* ,1993, 30(4), pp. 467-477.
6. Kotler, P., *Marketing Management, the Millennium ed.* Upper Saddle River, NJ: Prentice-Hill,2000.
7. Liu, C., Arnett, K.P. , Exploring the factors associated with Web site success in the context of electronic commerce, *Information Management* , 2000, 38(1), pp.23-33.
8. Min, Song & Keebler, An Internet-mediated market orientation (IMO) : Building a theory, *Journal of Marketing Theory and Practice*, 2002, 10(2),ABI/INFORM Global, pp. 1-11.
9. Narver J. C. and Slater J. S., The Effect of a Market Orientation on Business Profitability, *Journal of Marketing*, 1990,54(4), pp.20-35.
10. Sanzo, M. J., M. L. Santos, R. Vazquez and L. I. Alvarez, The effect of Market Orientation on Buyer-Seller Relationship Satisfaction, *Industrial Marketing Manamement*, 2003,

32(4),pp.327-345.

11. Srinivasan, V., Anderson, R., & Ponnavaiah, K., Customer loyalty in e-commerce: an exploration of its antecedents and consequences, *Journal of Retailing*, 2002, 78, pp.41-50.
12. Szymanski, D. M., & Hise, T. R., E-Satisfaction: An Initial Examination, *Journal of Retailing*, 2000,76(3), pp. 309-322.

GLOBALIZATION AND THE MULTINATIONAL FIRM

Jolene Marie Bon-Jorno, Sam Houston State University

Bala Maniam, Sam Houston State University, maniam@shsu.edu

Kathy Hill, Sam Houston State University, khill@shsu.edu

ABSTRACT

Globalization comes in many forms, including; economic, military, environmental, and social globalism and it can be measured through trade flows, factor movement, and price convergence. Globalization is mainly spread through the influence of multinational enterprises. The firms began as regional trading companies and have grown to become integrated networks of firms and affiliates, connected through a unified process and business mission.

CAN SMALL BUSINESSES COMPETE IN TODAY'S GLOBALIZED ECONOMY?

Marla Slaughter, Sam Houston State University

Bala Maniam, Sam Houston State University

maniam@shsu.edu

ABSTRACT

Today's globalized economy has developed through many economic and technological movements working together to integrate the universal economy. This paper discusses how the market environment created by a globalized economy affects small businesses and their ability to compete. The information assimilated regarding globalization and its' impact on small businesses has generated a wide array of opinions ranging from declaring substantial benefits to proclaiming a death sentence for such businesses. Some of these studies will be reviewed and explored to help determine the future of small businesses in an ever increasing global economy and the steps these entities can take to improve their chances of success.

In addition, the various aspects that augment globalization and how they each affect small business, as well as the history of globalization and how small businesses entered the international market place, will be reviewed. This study will also encompass governments' role in placing burdens on small businesses. By examining a wide array of issues affecting small businesses, a conclusion will be drawn with regard to their ability to compete in today's globalized economy.

Wojciech Nasierowski
Faculty of Business Administration
University of New Brunswick
Fredericton, NB, E3B 5A3 Canada
and Academy of Humanities and Economics, Lodz, Poland
e-mail: nasierow@unb.ca,
tel. 1-506-458 7338, Fax 1-506 453 3561

Some Observations Regarding Composite Indexes of Economic Performance

Abstract

This paper examines the information content of selected composite indexes, namely the Growth Competitiveness Index, the Knowledge Economy Index, the Human Development Index, the World Competitiveness Report Index and the European Innovation Scoreboard. These indexes are examined from the viewpoint of country rankings. It is argued that these indexes provide highly similar information, which brings to question the usefulness of such a variety of approaches. This paper also questions whether these indexes can adequately serve as policy setting mechanisms.

1. Introduction

Composite indexes are used in a variety of economic performance and policy areas. Such indexes integrate large amounts of information into easily understood formats and can be manipulated to produce desired outcomes. Despite this, there are several methodological problems regarding the creation of composite indexes (Saisana & Tarantola, 2002). Several issues arise when examining the accuracy and reliability of these indexes. Problems of missing data are imminent, along with the question of index sensitivity to the weighing of indicators and their aggregation (Freudenberg, 2003, p. 5). Composite indexes measure complex, dynamic systems, and “many of their properties emerge from interactions among the entities in them” (Katz, 2006, p. 893). It has been noted that problems arise when using these indicators due to the conceptual quandary between allocative efficiency (“are we doing the right things?”) versus technical efficiency (“are we doing things the right way?”). Further issues stem from problematic definition types, and the various taxonomies used to measure the consequences of achieved output (e.g., Seng Tan, 2004). Albeit these problems, there has been a noticeable proliferation of composite indexes (e.g., Archibugi & Coco, 2005, pp. 179-181). With several notable problems, the need and/or usefulness of such diversity among indexes is questionable. Despite being created with different intentions, and using varying series’ for calculations, these composite indexes may actually produce similar results when raking countries (e.g., Mirchandani, 2005). This occurs irrespective of whether they intend to measure level of innovative capability, competitiveness, productivity, level or wealth, or standard of living. To that end, this paper will seek to examine the specific information content of selected composite indexes.

In general, composite indexes utilize a variety of indicators (data series) to measure phenomena identified by their authors. Freudenberg (2003) provides an extensive overview of composite indicators of country performance, examining groups related to economy, environment, globalization, society, and innovation/technology. The following indexes have been selected for the purpose of this report: the Growth Competitiveness Index (GCI, 2005),

the Knowledge Economy Index (KAM, 2005; Chen & Dahlman, 2005), the Human Development Index (HDI, 2005), the World Competitiveness Report Index (WCY, 2006), and the European Innovation Scoreboard (EIS, 2005; EIS, 2006).

Composite indexes indicate which items of economic performance may contribute to the enhancement of an economy. This information may be useful in providing policy formulation suggestions for governments. An inherent assumption is made that some 'policies', as measured by these indicators, will produce similar results irrespective of specific and varying country context. This assumption, however, may be incorrect. An overview of literature results suggests that indexes used to measure economic phenomena demonstrate substantial overlaps and redundancy. From the viewpoint of outcome or statistical-type assessments, such variety may be deemed unnecessary. To that end, it is worth examining whether a more simplistic composite index of innovativeness can be formed, and eventually enhanced for a broader range of countries.

The primary purpose of this paper is to discuss the information content of indexes for each country. To that effect, the paper is organized as follows. Section two introduces key composite indexes of economic performance selected for examination. In section three, rankings produced by these related composite indexes are compared. A conclusion follows along with recommendations for further study.

2. AN OVERVIEW OF COMPOSITE INDEXES

1. The Growth Competitiveness Index (GCI) as published in the Global Competitiveness Report (2002) is based on three "pillars" (McArthur & Sachs, 2002):

- macroeconomic environment – assesses the stability of macroeconomic situations (e.g., through a tight monetary policy, low inflation) (macroeconomic stability, government waste, and country credit rating sub indexes);
- institutions – assure a favorable climate for long-term economic and business activities (e.g., contracts, law, and corruption sub indexes); and,
- technology – captures features of technological progress (innovation, ICT, and technology transfer sub-indexes).

These three sub-indexes are then combined to calculate the overall GCI.

2. The Knowledge Economy Index (KEI) of the World Bank is prepared based on the Knowledge Assessment Methodology (KAM) (e.g., Chen & Dahlman, 2005; KAM, 2006), and highlights the importance of knowledge for long-term economic growth. Some 80 data series are grouped as follows:

- performance indicators
- economic incentives
- institutional regime
- innovation system
- education and human resources
- information infrastructure.

This information forms the basis for identifying four sub-indexes, emphasizing the use of existing and new knowledge and the flourishing of entrepreneurship;

- an educated and skilled population to create, share, and use knowledge well
- a dynamic information infrastructure to facilitate the effective communication, dissemination, and processing of information

- an efficient innovation system of firms, research centers, universities, consultants and other organizations to tap into a growing stock of global knowledge, to assimilate and adapt it to local needs, and to create new technology.

3. The Human Development Index (HDI) is prepared within the United Nations Development Program, and is a “composite index measuring average achievement in three basic dimensions of human development – a long and healthy life, knowledge and decent standard of living” (HDI, 2005, p.271 and Technical note 1).

4. The World Competitiveness Yearbook Index (WCYI) is prepared by **IMD**. Its objective is to analyze the facts and policies that shape a nation's ability to create and maintain an environment of value creation for its enterprises and more prosperity for its people (WCY, 2006, p.15). Based on 312 criterion, including data from various sources and an annual Executive Opinion Survey, there are four categories:

- economic performance, which measures the macro-economics of the domestic economy;
- government efficiency, which evaluates the extent to which government policies are conducive to competitiveness;
- business efficiency, which assesses the extent to which enterprises are performing in an innovative, profitable, and responsible manner;
- infrastructure which denotes the extent to which basic, technological, scientific and human resources meet the needs of business.

5. The European Innovation Scoreboard (EIS) is published by the European Commission (Sajeva et.al, 2005; European Innovation ... , 2005). To create the EIS Index, 52 indicators are selected arbitrarily based on relevancy and data availability. Later they are grouped into five scales (categories) based on “expert knowledge”. For each scale principal component analyses have been performed to eliminate redundant information-wise items. Finally, the following five scales with 26 indicators have been identified (Items of these scales are listed in Sajeva et.al., 2005, pp.8-9):

1. Innovation drivers
2. Knowledge creation
3. Innovation and Entrepreneurship
4. Application
5. Intellectual property

Each composite index consists of sub-indexes, where all items are equally weighted. This methodology may be challenged for correctness in terms of selecting (and grouping) indicators/ Despite this numbers should indicate the number and nature of the factors that describe the idea. Moreover, several items are highly correlated – they carry the same information with regard to statistical significance of results (and country rankings). This methodology, in principle, is characteristic of all composite indexes presented.

3. INFORMATION CONTENT IN COMPOSITE INDEXES

In light of the previous discussion this section will revert back to the opening research question – what is the information content of indexes?

As illustrated in the earlier section, there are several composite indexes. They use somewhat similar methodologies and data series to calculate a composite index. To that end they may produce similar results – irrespective to whether they intend to measure the level of

innovative capability (EIS), level of competitiveness (WCY), level of well-being (HDI), level of growth capability (GCI), or knowledge base (KEI).

To assess the magnitude of information content in various indexes the Spearman correlation coefficients (e.g. Siegel and Castellan, 1988) of country rankings are computed (Table 1). These are obtained from any two indexes and test whether the resulting correlations differ statistically from zero. Additionally, computations of two easily accessible indicators have been included, that may have similar information content as composite indexes: namely Purchasing Power Parity (PPP), and Labor Productivity (LP).

Table 1. Spearman correlation coefficient for various indexes (upper triangle – p value, lower triangle correlation coefficient and N)

	GCI	KEI	HDI	WCY	EIS	PPP	PROD
GCI		.000	.000	.000	.000	.000	.009
KEI	-.944		.000	.000	.000	.000	.002
	31						
HDI	-.777	.850		.000	.000	.000	.000
	32	32					
WCY	-.867	.834	.805		.000	.000	.001
	29	29	29				
EIS	-.759	.880	.781	.649		.000	.002
	32	32	33	29			
PPP	-.687	.720	.900	.837	.626		.000
	28	28	28	28	28		
PROD	-.484	.563	.779	.588	.873	.873	
	28	28	28	28	28	28	

There is overwhelming evidence that the ability of one index to provide additional information over another is questionable at best. The p-values of these non-parametric Spearman correlation tests are all under 0.001. Therefore, the hypothesis that different indexes produce different country rankings can be rejected. Statistical data suggests that composite indexes carry similar information. To that end whether we examine the concept of innovativeness or wealth, we are essentially discussing closely related aspects of economic prosperity.

As was indicated previously composite indexes consist of sub-indexes. Thus, it is interesting to note, whether or not these sub-indexes are highly correlated, and whether they create yet more redundancy. Indeed, the Spearman correlation coefficients indicate that sub-indexes within each composite index are highly correlated, and almost exclusively fall below the 0.001 level. This indicates that not only composite indexes, but also sub-indexes are redundant from for the purpose of ranking countries. One can easily formulate the list of common topics used to assess the level of a country's general prosperity. These will include two broad sets of indicators:

- those related to the macro-economic situation and business arrangements such as PPP; labor productivity; macroeconomic environment and incentives, institutional/governmental regime; business efficiency; and standards of living; and,

- those related to knowledge/education; information/technology/communication infrastructure; and innovation systems.

4. SUMMARY

It should be noted that the preceding discussion does not negate the usefulness of these indexes. Rather, the results question the need for yet other indexes that replicates the ranking of alternate or better established measures.

Each index captures some information related to economic improvement. Since these items are somewhat correlated, it should be asked which acts as stimuli for the development of other ideas. For example, numerous studies (e.g. McArthur & Sachs, 2002, Porter, 1990, Rutten & Boekema, 2005; TEP, 1992; Fagerberg, 1994; Sterlacchini, 2006) highlight the importance of innovation to economic development and well-being. Yet, it remains unclear whether countries are innovative because they are rich, or are rich because they are innovative. Similar questions may be asked regarding other concepts discussed by selected composite indexes.

To be clear, it is at times suggested that composite indexes may serve to guide policy settings. However, data series used in these composite indexes change almost every year. Consequently, there is no possibility of identifying whether policy changes have contributed to the improvement of desired operational outcomes.

Acknowledgement: Support to this study has been granted by the University of New Brunswick, Faculty of Business Administration and is gratefully appreciated.

SOURCES

Archibugi D., Coco A., 2005, Measuring technological capabilities at the country level: A survey and a menu for choices, *Research Policy*, Vol 34, pp.175-194.

Arundel A., Hollander H., 2006, 2006 Trend Chart Methodology Report: Searching the forest for the trees: "Missing indicators of innovation, Maastricht Economic Research Institute on Innovation and Technology.

Chen D.H.C., Dahlman C.J., 2005, The knowledge Economy, the KAM methodology and World Bank operations, the World Bank, Washington DC 20433

EIS 2005, European Innovation Scoreboard 2005: Comparative Analysis of Innovation Performance, European Trend Chart on Innovation.

Fagerberg J., 1994, Technology and international differences in growth rate, *Economics of Innovation and New Technology*, Vol 3, pp183-200.

Freudenberg M., 2003, Composite indicators of country performance: A critical assessment, STI Working paper 2003/16, Industry Issues, OECD JT00153477

(GCI) The Growth Competitiveness Report, 2004-2005, 2005, Global Competitiveness Report of the World Economic Forum.

(GCR) The Global Competitiveness Report 2001–2002, 2002, Global Competitiveness Report of the World Economic Forum.

(HDR) Human Development Report: International comparison at a crossroads: Aid, trade and security in an unequal world, 2005, Published for the United Nations Development Program, New York.

(KAM) The Knowledge Assessment Methodology and Scorecards 2006 - <http://vle.worldbank.org/gdln/Programs/kam2001/methodology.htm> - assessed on 24/08/2006

Katz S., 2006, Indicators for complex innovation systems, *Research Policy*, Vol 35, Issue 7, pp.893-909.

McArthur, J. W., Sachs J.D., 2003, The Growth Competitiveness Index: Measuring Technological Advancement and the Stages of Development, in the Global Competitiveness Report 2001–2002. New York: Oxford University Press for the World Economic Forum.

Mirchandani D., Economic and social indicators of global competitiveness: An analysis of country rankings, 8th International Eastern Academy of Management Conference, (CD Rom version), Prague (June 1999).

Porter M., 1990, Competitive Advantage of Nations, *Harvard Business Review*, Volume 90, pp. 73-92

Rutten R., Boekema F., 2005, innovation, policy and economic growth: Theory and cases, *European Planning Studies*, Vol 13, issue 8, pp:1131-1136.

Saisana M., Tarantola S., 2002, State-of-the-art report on current methodologies and practices for composite indicator development, Joint Research Centre – European Commission, I-21020 Ispra (VA) Italy.

Sajeva, M., Gatelli, D., Tarantola, S. and Hollanders, H. (2005), Methodology Report on the European Innovation Scoreboard 2005, European Commission, Enterprise Directorate-General.

Seng Tan, B., 2004, “The Consequences of Innovation” in *The Innovation Journal: The Public Sector Innovation Journal*, Volume 9(3), pp.1-42.

Siegel, S. and Castellan, N.J. (1988), *Nonparametric Statistics for the Behavioural Sciences*. McGraw-Hill, New York.

Sterlacchini A. 2006, Innovation, knowledge and regional economic performance: Regularities and differences in the EU, a paper presented at the DRUID Summer Conference 2006 (accessed via google on Sept 10, 2006)

TEP, 1992, Technology and the economy: The key relationships, The Technology/Economy Programme, OECD, Paris.

(WCY) World Competitiveness Yearbook (2006) IMD, Lausanne, Switzerland

Oh, you are an Artist - then you are Involved in International Business

Jan Bodin, Ph.D.

ICN Business School,
13 Rue Michel NEY
54000 NANCY, FRANCE
Telephone: +33.(0)6.32.01.20.84
Facsimile: +33.(0)3.83.39.64.80
E-mail: Jan.Bodin@icn-groupe.fr

Submitted to:
The 2008 Northeast Decision Science Institute (NESDI) conference,
Brooklyn, New York, March 28-30.

January 23, 2008

Abstract

This paper reports the result of a project involved in investigating the perception of business and international business among artists. Ten interviews among a diverse group of artists from different countries were carried out during 2007. All of the artists are or were fellows at Academie Schloss Solitude in Germany. An exploratory research approach gave interesting findings regarding the artists' way of perceiving their profession, their attitude towards business and business activities within their field. This also has consequences for how they defined themselves and what is needed in order to be established as a 'real' artist.

It is obvious that artists today are a group that can be considered running a small project based company. They need to finance their activities, distinguish the audience (customers), decide upon which project to carry out etc. Since today's artists also are travelling in different countries to learn and explore their profession, it covers many interesting consequences connected to the field of international business. Among this small group differences could be established that might be linked to their cultural background. A special focus on their approach to fund their projects, and their (in some cases) inability to see the potential future (international) market of a product in a project is made.

Introduction

This paper addresses the topic that young and upcoming artists are in many ways behaving similar to small born global firms. The study is based on ten interviews with artists with titles or "labels" like; photographer, writers, painters, architect/designer, musician, reality engineer, and play writer. The artists also represent a wide range of nationalities covering countries like Germany, Indonesia, Japan, South Africa, Sweden, UK, and USA.

The interviews were carried out during 2007 at Akademie Schloss Solitude in Germany where all respondents were resident fellows, i.e., were working and living on a full time grant by the academy. Schloss Solitude is Germany's largest artistic academy funding approximately over 60 full time scholars from all over the world each 2-year period. Each interview lasted between just over 30 minutes for the shortest to 2.5 hours for the longest. Each interview was recorded in digital format to prevent any loss of information in addition to taking notes. The respondents have also been given the opportunity to comment on the text.

Artists, Business & Born Global Firms

One of the respondents addressed the importance of defining when you can label yourself a true artist. According to the respondent, a keystone to label someone an artist is that their main time is spent working with their profession. Someone that is working, for example, in a factory during daytime, and is working as a writer in the evenings or during the weekends does not qualify according to this definition. Another possible definition is based on what the individuals' main income is derived from. According to these definitions, the respondents fit both categories. They are all fully funded and working full time with their profession.

One of the painters stressed the importance of keeping his creative freedom and commented that he did not perceive success the same way as an Asian colleague he once visited that was highly popular, but painted "on demand". This theme reoccurred in different variations with many of the respondents. It seems highly important for them to keep their "creative freedom".

I.e., they are not really interested in creating structures and standardized behaviours, but prefer to uphold a high degree of flexibility. This behaviour/myth of the struggling and starving artist that is willing to undertake severe problems in order to explore their creative freedom has a long tradition. During the late 1980's to mid 1990's some art schools started to implement courses like "Survival skills for Artists" in order to at least partially change their perception of their profession. They realized that it was not enough to just teach art since their students after they graduated needed to interact with numerous business representatives and in a sense become a business person themselves in order to be successful (Metaxatos 1996).

The idea that most firms expand their business incrementally, like rings on the water has been the traditional view on internationalization (see for example Bilkey and Tesar, 1977). In the beginning of the 1990's new research pinpointed that a different behaviour was evident in certain companies and the term international new ventures and born global was emerging (see for example (Rennie 1993) and (Aspelund et al. 2007)). Among "born global" firms it has been established that the profile of the owner/manager with regard to international entrepreneurial orientation, geocentric mindset, prior international experience, and learning orientation is positively related to their marketing capability (Weerawardena et al. 2007).

Funding & International Movement

Similar to upstart theories of a company (see for example Churchill and Lewis, 1983) there are a certain period where external financing is needed. A normal company searches for venture capital via private investors or banks to survive over this period. All respondents in this study have reached a level of reputation within their profession that made it possible for them to be handpicked by the awarding committee at the academy. Some of the artists expressed the view that they had a period of some years when they had the opportunity to apply to (and hopefully receive) grants from different academies around the world. That the different academies were located in countries like for example Germany, Spain, Holland, and USA was not seen as an obstacle at all. With few exceptions, they all saw travelling to different countries rather as a way of gaining more experiences, time to create, and naturally receive funding than as an obstacle. This is an important way of funding for this category of artists as a way of bridging their capital need during their "start-up" phase.

Another option artist usually signs with and agent, broker, or some other form of representative. This representative has an important role to play since he or she will act as the artist's manager/ mentor as well as the distributor of the products the artist is producing. During the interviews it became evident that the commission that the agent/broker charges differs depending on artistic focus. Starting around 15% for commissioned work like translations and editing, all the way up to 50% for agents and galleries.

Interesting to note is that most of the artists have the possibility to carry out smaller projects on the side that are partially connected to their profession as a source of extra income. A writer can do editing or translations assignments, while a photographer can take on assignments to cover a certain event like a wedding or a show. Graphic oriented artists on the other hand can work for companies with design- and logo issues. When talking to the respondents most of them perceived these side projects as a bread and butter assignments that they would stop doing as soon as their financial situation becomes better.

Wide Range of Products

An artist's product can take a multitude of shapes and structures. From the more traditional types of output like a book or a painting to put on a wall, to different types of events like a play, concert or performance. In that respect it is possible to say that the products can span from tangible to intangible. Interesting to note is that even if the core product is an event that can be compared to what Grönroos (1998) label the core product/service a play-writer can (besides setting up the play) sell written copies of the play, programs, etc. as tangible "add-ons". A designer of an event can similarly sell sketches of preliminary versions etc. The variety of projects that are carried out, both in scope and in time is fascinating. Besides the more "normal" projects with output like paintings, novels, etc. there are others that sticks out. Some examples that illustrates the business linkages follows below:

One of the artists has been travelling to many different countries, setting up projects with a clear business character. One project involved setting up an Esperanto lunch restaurant selling only boiled eggs for a shorter period. Even if the main focus was on expressing communication though Esperanto the setup was a real business activity. A similar project, an Esperanto restaurant/hot dog stand, was also established on Island.

The projects can be initiated as a visionary idea and be expanding over time. The photographer had a visionary idea of documenting "contemporary Europe". He started on the project without being fully funded, but is still planning to carry on his project over a period of six to eight years. This commitment will force him to travel all over Europe to photograph different places. In order to accomplish that, he is actively searching for funding in many different countries. So far partial funding/sponsorships has been found in USA, Germany and Austria, but contacts/applications have also been made with companies and funding agencies in for example Holland, Spain, and Suisse.

Another interesting example is one of the artists that had carried out a project where she combined music with movement of skateboards. By connecting the skateboard to a computer, the person on the skateboard could create his or her own music on a computer. While carrying out the final testing of the skateboards before the performance she was approached by people that asked when it was possible to buy this product in a store. Even with this obvious possibility to investigate market opportunities based on prospective customers, she did not proceed with the project. She explained that she was focused on the artistic side and to explore the possibilities. After reflecting on the issue she said that she would be happy if she could find someone else to work with her on exploring the business side of her ideas, but that she did not find it personally interesting.

Conclusions

First, the group of respondents in this study shows a great variety in nationalities and how they express their artistry. On the other hand, the group is homogeneous with regard to the fact that they have all been accepted as scholars at the same academy. The fact that a large well established German academy is recruiting their scholars from all over the world is a clear indication that even artists follow the constantly stronger trend of internationalization and globalization that affects all parts of today's society.

Second, we need to start realize that even if the driving-force of today's artists are still idealistic and visionary, the harsh reality is that being an contemporary artist can be compared to managing a small firm. In some cases the artist can distance themselves from the business

side of their firm by working with some sort of agent or representative, but they still need to manage their own activities. The respondents have all be acting in a deliberate way and used or at least reflected on the international arena as a tool for their career. In that sense they fit the born-global category even if it is difficult to benefit from categorizing them for example into labels like the Respondent, the Opportunist, the Experimentalist, and the Strategist mentioned by (Freeman and Cavusgil 2007). As been shown with the example of the photographer it can involve planning, networking, and negotiations with companies and organizations from many different countries.

Third, even if the reality is that many artists can be compared to a manager of a small firm, it does not automatically mean that the artists reflects on this, or want to explore this road including take advantage of upcoming business opportunities. As the skateboard project illustrates, even with a clear request from a potential buyer, the artist was so focused on the performance and outcome of her artistic work that she excluded everything else.

Forth, there seems to be only partial interest among the artists to seriously explore emerging business opportunities. It might therefore be suitable for larger institutions and funding agencies focusing on this group to provide support and help in these issues. Politicians and business leaders all over the world are often arguing the constant need for more new ideas and innovations to explore – here is a chance to unleash the potential of a neglected group up to this point.

Fifth, there might be cultural reasons to the respondents' different behaviour. Due to the small group of respondents it is not possible to establish if differences depends of individual characteristics or cultural/educational upbringing. Although, it was noticeable that for example the two Americans in this study were able to express themselves in more defined business terms compared to the others, and that they seemed to have analyzed their potential market and even considered and implemented some CRM-techniques.

References

- Aspelund, Arild, Tage Koed Madsen, and Øystein Moen (2007), "A review of the foundation, international marketing strategies, and performance of international new ventures," *European Journal of Marketing*, 41 (11/12), 1423-48.
- Bilkey, W. J. and George Tesar (1977), "The Export Behavior of Smaller-sized Wisconsin Manufacturing Firms," *Journal of International Business Studies* (8), 93-96.
- Churchill, Neil C. and Virginia L. Lewis (1983), "The five stages of small business growth," *Harvard Business Review*, 61 (3), 30-50.
- Freeman, Susan and S. Tamer Cavusgil (2007), "Toward a Typology of Commitment States Among Managers of Born-Global Firms: A Study of Accelerated Internationalization," *Journal of International Marketing*, 15 (4), 1-40.
- Grönroos, Christian (1998), "Marketing Services: the case of the missing product," *Journal of Business and Industrial Marketing*, 13 (4/5), 322-38.
- Metaxatos, Despina (1996), "Teaching artists how to succeed in business," in *American Artist* Vol. 60.
- Rennie, M (1993), "Global competitiveness: born global," *McKinsey Quarterly* (4), 45-52.
- Weerawardena, Jay, Gillian Sullivan Mort, Peter W. Liesch, and Gary Knight (2007), "Conceptualizing accelerated internationalization in the born global firm: A dynamic capabilities perspective," *Journal of World Business*, 42, 294-306.

USING INNOVATIVE IDEAS TO DEVELOP CREATIVE WINNING STRATEGIES

ABSTRACT

Building on (and continuing) the structured analysis of the present market and of possible future market trends, and on an analysis of a specific company's strengths and weaknesses and possible strategic moves which might affect its performance in that marketplace, this paper describes winning strategies other specific companies have developed in response to the observed and anticipated driving forces in competitive markets. The studies are designed to stimulate further thinking about ways to brainstorm new ideas and write new scenarios for possible alternatives a company under study might want to explore.

THE TRADITIONAL SYSTEMATIC ANALYSIS: STUDYING THE EXISTING SITUATION

Since a manager in a specific company formulates strategies within a specific competitive market, the strategic management process in practice requires studying both the kind of company under study and the industry/competitive market segments involved, as well as their impact on winning. This is basically and initially the traditional systematic (structured) competitive market analysis described in many books [for example, Mockler 2002(A)]. This base building phase is especially important for those unfamiliar with the market under study or for those with little business experience or for those dealing with multinational markets. While this phase is less information-acquisition intensive for an executive with experience in the market under study, the detailed re-conceiving and reconfiguring of market and company information required to stimulate winning strategy development, as well as the inferential reasoning required, can be challenging.

THE SEARCH FOR NEW TRENDS: THE INTERMEDIATE AND LONGER-TERM

Business Process Value Chain Analysis of the Existing Market

A commonly used approach to studying the trends which might affect the intermediate and long-term future is to examine each component of the existing value chain, as well as the value chain as a whole, in order to identify any emerging shifting patterns. This kind of analysis requires conceptual, as well as associative and inferential reasoning, cognitive skills at times different from the skills needed for the systematic traditional analysis described in the preceding section. It is an approach often used in brainstorming sessions to stimulate creative, innovative thinking about new trends and new strategies to exploit them.

Overall, Especially Outsourcing. Many companies have moved away from having integrated internal value chains, that is, performing all the functions in the business process value chain themselves. One major clue used to identify where this is happening is found in areas where outsourcing is increasing. For example, in the chip manufacturing industry several companies began focusing on just the design and selling of chips, that is becoming so-called "Fabless" companies, which left the fabrication of chips to outside foundries, a major new opportunity area exploited by many new and existing companies, especially chip foundries located in Taiwan [Cane 2002; Chang and Mockler 2000]. So great was the opportunity that IBM actually built its own foundry in Armonk, New York, a third of which's production will serve IBM chip needs and the rest make chips for other "Fabless" companies [Lohr 2003(A)].

In early 2002, outsourcing of noncore business processes, such as travel services, insurance administration, employee benefits, and collections, was reportedly growing both in the U.S. and Europe [Maitland 2002; Roberts and Luce 2003]. The question to answer is, in your industry are there components in the value chain which can be isolated and exploited for profit or cost savings and sold off or outsourced, or are there opportunities for reintegrating components?

Suppliers and Customers. The shifting strengths of suppliers and customers, which were described in 1980 in Michael Porter's classic market study *Competitive Strategy* [Porter 1980], can also be harbingers of future trends. For example, as the number of media and broadcast channels expanded with the growth of cable systems, the major networks found themselves in a profit squeeze as program content suppliers grew stronger and increased their prices through competitive bidding among the major networks. At the same time, customers had many more viewing options--over one hundred viewing channels on some cable systems, for instance. Broadcast companies caught in this kind of price squeeze have the option of expanding supply sources, for example by developing their own programming content (often through buying content companies), or the option of entering other market segments, for example by buying or creating new broadcast channels (news channels, for instance). Content companies in turn were also expanding into the delivery area by buying broadcasting, cable, and satellite companies [Lowry, et al. 2004].

Manufacturing/Production. With the advances in technology and increasing customer sophistication, it is becoming more feasible to profitably serve smaller and smaller customer segments, to the point of mass

customization, often referred to as BTO (build-to-order) [Marsh 2001], which opens up a whole window of new opportunities and keys to success.

The customization decision involves determining where, if anywhere, the potential exists for serving smaller segmented markets, should they be developing. It can involve even going so far as to serve individual customers in a customized way, something done by National Bicycle in Japan which has a computerized manufacturing system which enables it to make custom ordered bicycles for about 15% more than top-of-the-line mass-produced bicycles and deliver them in less than two weeks [Westbrook 2000]. On the other side of the globe, an Italian shoe company had retail outlets take customer foot measurements and then transmit them by computer to its Italian factory where shoes are made to order and delivered in about two weeks. In its U.S. outlets, the price for a pair of men's shoes from this Italian manufacturer averaged about \$140, comparable to the cost of high-quality, mass-produced shoes but well below the \$500 or more that a hand-made pair usually costs. Lands' End now successfully sells custom-made pants over its Web site [Tedeschi 2002].

One enabling approach in this value-chain segment is to use a common product platform (as do Dell Computer and many auto manufacturers, especially the Japanese who are far ahead of the U.S. carmakers in this area [Shirouzu 2003; Welch and Kerwin 2004] and then add prefigured custom modules at a late stage of the manufacturing process [Editorial Staff 2001].

Sales and Distribution. Patterns of change can also be studied in the sales and distribution area. Disintermediation has occurred in many fields, for example, as the Internet has enabled selling directly to consumers and eliminating the middleman or salesman, often as in the stock brokerage industry doing this at much lower cost. This trend, generated by available technology and changing customer buying patterns, has been a trend hurting many wholesalers and retailers, who have had to take steps to compete online.

Advertising/Promotion. In the media field, the big-three television networks during the 1990s found that advertisers would continue to pay ad-rate increases even as many viewers defected to cable channels, an emerging trend which reversed past pricing patterns where ad-rates were directly related to the size and income level of the market. Subsequently, the lessons learned in television were applied to the newspaper field, where the same seemingly illogical and contra productive non-traditional advertising pricing strategy was successfully used.

Products. In the product area, several types of pattern shifts can be identified, including: from product-to-product pyramid, product-to-customer solutions, and product to brand. This analysis is designed to spot trends that are occurring early enough to be able take steps to exploit them or limit damage to your existing business.

Product pyramids build product lines spanning high to low end prices/products, with different styles, designs, and performance. Increasing consumer sophistication and income stratification can create situation circumstances which create opportunities for these kinds of patterns. Swatch successfully built a higher priced line on top of its initial low-priced lines, for example. Some computer manufacturers, such as Microsystems, were slow to respond to an obvious consumer trend towards much cheaper basic computers, as well as smaller portable ones, and so lost market share by not having a broad enough product line [Morrison 2003; Tam 2003].

Customers often now look for more than just products; they seek broader solutions to their marketing needs. Intel dominated its market and far outdistanced its major competitor, Motorola, by providing total customer solution -- for example by helping customers such as Ford design products using its chips.

In 2000 Microsoft made a daring strategic move designed to "transform Microsoft from a purveyor of packaged software increasingly into a supplier of information-age services like digital wallets, software subscriptions and e-commerce portals" [Brick 2001; Greene 2001]. In 2003, it again made major moves to reshape the company to meet changing market needs, with products and services such as software that powers mobile phones, online game machines, television services, and Business-To Business (B2B) internet buying-selling services [Lohr 2003(B)]. These strategic moves began with careful analyses of customer needs and their economics which led to identifying ways in which a company can provide unique products and services solutions that can improve its customers' systems, service, and profitability.

Branding. Frank Purdue was able to dominate the chicken market, an almost commodity type product, by both taking steps to insure quality and then heavily promoting his brand name with tangible evidence such as its golden color. Sara Lee also was able to dominate its markets with a highly-focused brand approach. These firms built brand image though multiple channels. Virgin Airways went beyond product brand development and developed a corporate brand image [Towers 2001] which it used as a profit multiplier, enabling expansion across many product lines well beyond the airline industry.

The risks involved in basing a brand image on a person was evident in 2002 when Martha Stewart, whose name is linked to all her products and ventures, was accused of illegal stock trading, a major problem since the success of her ventures depends in large measure on a squeaky clean image [Beard 2003; Rose 2002].

In general, branding is one major way to create distinctive winning strategies for a company and can be done either through creative promotion or through developing distinctive products, as seen from the above examples. At the same time, more recently in-store brand products have been growing in popularity and competing effectively with national brands [Boyle 2003].

The Impact of the Internet

The preparatory analytical phases described above involve decomposing, that is, analyzing, the new driving forces (identifying components) and inferring their impacts on the market (the external analysis). Using a slightly different variation of that approach, Hamel and Sampler [1998] did a useful general analysis of the characteristics of e-commerce and ways it can impact industry structures, including ways in which firms can exploit and will be affected by e-commerce in general. The study also shows the ways in which the Internet has enabled meeting current customer desires for products (what, where, when, how, and at the price they want them).

Going further, in the *Financial Times* on December 30, 2000 the late Peter Martin gave the following view of the Internet's future based on his market studies [2000, p. 9]:

"What will happen to internet companies? There will be no internet companies. That is, by the end of 2001 the integration of the 'real' and 'virtual' economies will be complete. Unsuccessful dotcoms will have vanished, successful ones will have amalgamated with their opposite numbers in the conventional business world. And traditional businesses will have added an electronic dimension to their activities. Internet technology will be maturing, as innovation slows and conventional software engineering disciplines infiltrate the sector. The arrival of broadband connections and enhanced handheld devices will produce a flurry of new pricing models, in which users are asked to carry a greater share of the cost of producing and publishing content -- but most will fail. The shift of the Web's center of gravity away from US sites and the English language, long predicted, will be clearly evident in 2001. This is just one way in which the web will no longer be seen as an agent of revolutionary social and economic change but as an enhancer of existing trends."

Only time will reveal the accuracy of these predictions. They can, however, be useful in stimulating future thinking and further investigative analysis. For example, by early 2001 thousands of companies such as Western Union, Encyclopedia Britannica, Nabisco, United Airlines, Deutsche Bank, and Ford, were all incorporating or struggling to incorporate internet technologies into their existing businesses [Hansell 2001; Ridgway 2000]. For instance, admitting its two-year-old online banking venture was a failure, in mid-2001, Bank One folded its independent online WingspanBank.com into the rest of its online banking operations [Coleman 2001].

In another situation, Charles Schwab & Company, a discount financial brokerage firm, determined through an analysis of the internet capabilities, such as the one Hamel and Sampler [1998] did, that a pattern of the growing use of the Internet by individuals in many businesses might be extended to the brokerage business. Following up on this pattern analysis and associating that pattern with the way the discount brokerage business worked suggested an opportunity to further reduce trading costs for customers and make ordering easier. This was initially very successful for Schwab, but later the company temporarily encountered major problems [Editors 2003; Gasparino and Brown 2001].

Viacom, a large media conglomerate, also took a **disciplined integrative evolutionary approach**, which used the Internet in many different ways to support and extend its existing operations such as MTV, Nickelodeon, CBS, its radio stations, Paramount Pictures and Blockbuster retail stores -- a sort of brand extension strategy [Markoff 2004]. This was in contrast to the more extensive transformation Internet strategy used by Bertelsmann and Vivendi Universal, other media conglomerates, which led to their major problems in 2002 arising from trying to reshape existing media strategies to incorporate the Internet more fully [Johnson and Clow 2002; Landler and Fabrikant 2002].

Moving from the other direction, many dotcoms were expanding into traditional businesses. For example, several dotcoms which published internet magazines were starting companion print magazines [Coffey 2000] and e-retailers and other online businesses were making use of direct mail advertising and mailorder catalogs [Neuborne 2001; Tedeschi 2001].

In another situation, Barnes & Noble, a book retailing chain did not take internet booksellers seriously when Amazon.com started and later was required to scramble quickly to meet the real threats to its retail business posed by Internet selling (which essentially eliminated the retailing intermediary) and to exploit a new way of doing business, which integrated the Internet with its retail operation, a move which apparently was very successful. Its

major retail competitor, the Borders Group, was not so fortunate and closed its online operations in mid-2001 and, instead, let Amazon.com serve its customers [Hansell and Kirkpatrick 2001]. At the same time focused on innovative ways to improve its in-store operations, since the online book sellers, major discount stores such as Wal-Mart, and price clubs and book clubs were cutting deeply into traditional bookstore sales [Kirkpatrick 2003; Trachtenberg 2002].

Ignoring the Internet threat, then, can lead to costly failures. In another situation, Reed Elsevier bought the Official Airlines Guide, the voluminous printed airline timetable, for \$416 million (plus \$282 million in penalties for overcharges), only to find eight years later that the Guide had been replaced by internet sources and in 2001 carried a written-off value of zero when it was disposed of [O'Connor 2001].

During the early 2000's, a wide range of stand-alone internet companies, such as EToys, were going out of business [Beard 2001(A); Richtel 2001], while others were succeeding, such as FreshDirect, an online grocer [Lee 2003; Mullaney, et. al., 2003]. Ironically, one of the major internet success areas has been the peddling of smut and providing access to illegal activities. And in late 2002, MGM Mirage planned to be the first United States company to start an internet gambling casino. There were at that time 2000 such sites all operated overseas [Richtel 2002].

More recent studies have shown that other national economies, such as that in China, are also encountering similar experiences with the very mixed and often confusing impact of the internet [Smith 2001]. Such international and domestic general and individual company studies provide useful information about future industry/competitive markets, opportunities, and success characteristics, all of which would be incorporated into a thorough situation analysis.

The Impact of Technologies

Developments in other technologies can also be studied and tested for their possible future impact on business. For example, Wal-Mart, working with MIT in October 2001 and for a time with Gillette in 2003, was testing a new wireless tool (initially called "Radio Frequency ID Tags" or RFIDs) which can be printed or painted onto or attached to individual products (and/or by extension clothing or even skin) that would enable tracking the location and usage of a product through this radio/wireless device and which can be directly linked to computers [Booth-Thomas 2003; Ewalt 2003(A)]. These tags were also introduced in Germany in 2003 [Austen 2003]. In mid-2003, Wal-Mart announced it would begin requiring groups of suppliers to start using the Tags in 2005

[Murphy and Hayes 2003] although at the same time it shelved its RFID pilot program with Gillette [Ewalt 2003(B)].

For instance, such tags can substantially reduce in-store theft, enable producers to know just how long an individual consumer took to use a product, such as a tube of toothpaste, and generally produce a wide range of detailed market information [Deutsch and Feder 2003; Ewalt 2003A]. The tests were begun when the cost of this technology was predicted to drop from the present five cents to only one cent per individual product. The implications of this technology are enormous, extending beyond business to finding lost or stolen autos or helicopter parts (the government was developing a radio frequency system for this), tracking the location of children and other people, and the like. At the same time, all of this potential also raised many questions about invasion of privacy, since RFID tags can also be attached to an individual's skin [Konicki 2001; London 2003B].

The impact of digital technologies on products, services, and ways of doing business are found in many companies, for example at Nokia and Motorola (cellular phones), at local radio stations expanding nationally, and even at Microsoft (digital wallets). At times, new developments have occurred when breakthroughs in one area (for example, electrical engineering) are applied to an unexpected new area (for example, biotechnology and drug development) [Anders 2003]. Other similar studies done at the same time briefly reviewed the potential of other emerging technologies such as Biometrics, always-on wireless networks, privacy management, and banking software applications [Russell 2002; Schwartz 2002].

Studying Other General External Patterns Affecting The Longer-Term Future

Studies have identified many trends expected to dominate business over the longer-term future. For example, Wind and Main [1998] believe that marketplace emphasis is shifting from capital to knowledge, producer to consumer, Atlantic to Pacific, Japan to China, international trade to electronic commerce, computers to Internet, and money to people. As for organizations, emphasis appears to be shifting from mechanistic to organic, engineering to ecology, corporations to individuals and networks, horizontal and vertical to virtual integration, and business processes to cultures. As for individuals, emphasis is shifting from hard work to hyper effectiveness, faith to hope, security to uncertainty, current to future careers, and loyalty to courage.

Major technological developments are furthering these trends. Many of these can be disruptive or enabling depending on a manager's perspective and business situation [Skapinker and Griffith 2000]. For example, while the impact of the Internet has proved ubiquitous and major, its impact may be accelerated as mobile (wireless)

technology (called by some m-commerce) enables access to the Internet and related technologies anywhere and anytime [Cole 2000; van Grinsven 2000], though even here the nature and extent of the impact of m-commerce is uncertain and the potential return on the expected \$300 billion in mobile licenses and equipment in 2001 questionable [Beard 2001(B); Romero 2001]. For example, in mid 2001 sales of mobile phones were running almost 10% behind a year earlier [George 2001].

STUDYING THE COMPANY INVOLVED: THE INTERNAL ANALYSIS

During the situation analysis, studies are also needed of the investor or company involved, and of creative ways in which his/her/its resources can be effectively put to work in light of the new market forces anticipated, a process which often involves brainstorming sessions designed to generate new scenarios. This was the situation faced by Celera Genomics, as it struggled to figure out ways to commercially exploit the completion of the human genome sequence mapping [Pollack 2002]. A specific relationship study is needed to explore the ways the new force can be exploited strategically company-wide, at the strategic business unit level, and/or operationally (compressing the value chain, dramatically lowering transaction costs, and increasing customer service, for example). The final strategic model or approach will in some ways be unique to the firm, as well as the industry involved.

An indispensable step in this company analysis involves studying its competencies, limitations, and available resources, since this is the perspective which can significantly impact on strategy formulation and implementation. During this phase, a manager intuitively begins formulating preliminary ideas of possible alternative strategic directions for the firm under study.

For example, Microsoft's position in its industry and competitive market was unique in many ways and this uniqueness dictated its strategies. First, its Windows operating system was the industry standard, since it was used in 90% of personal computers (PCs) sold. This gave it a major degree of control over the use of competing PC software and had led to a U.S. Department of Justice lawsuit. Microsoft is an example of a large cash-rich leading edge firm (with a major product that had become an industry standard) in a rapidly changing high tech industry with a high degree of uncertainty which followed an unstructured approach to strategy formulation, an approach appropriate for the industry and company involved. It is an example of how strategies and strategy formulation are dictated in many ways by specific external and internal situation factors. It is also an example of how strategies and strategy formulation can vary from situation to situation.

Some observers have gone so far as to say that Microsoft is an example of a company with no strategy other than remaining dominant in its field [Brown and Eisenhardt 1998; Downes and Mui 1998]. This strategic perspective is evident in the strategic moves by Microsoft to block out all competitors to its Windows software, such as by going so far as establishing a fund to enable Microsoft to give its dominant Windows software free to governments and large institutions thinking about using Linux, a free software, as well as to having its salesmen lie about their identities at trade shows [Fuller 2003]. According to reports on its progress, Microsoft's strategy seems to be working to some degree, in spite of U.S. law suits attempting to contain its growth [Buckman 2002; Greene 2001].

It is clear that different businesses dictate different strategies because of the resources and business processes involved in succeeding in their situation -- for example, UPS in the air cargo/mail business has substantially different requirements from those needed to succeed at a multinational automobile company, such as Toyota. Even within the same industry, companies such as Toyota and Renault have different resources, business resources, and external requirements, such as union relations, government relations, and available facilities, products, and resources that can affect the kinds of strategies formulated.

FORMULATING AND IMPLEMENTING WINNING STRATEGIES

This paper provides many successful and unsuccessful examples of how such trends affect formulating and implementing winning strategies over the longer term. For example, Lou Gerstner, former CEO of IBM, developed an integrated longer-term **unarticulated** strategic vision for IBM, a vision in his case linked to the concept of network computing and electronic commerce, rather than one linked solely to personal computers. Within this vision, communications, rather than computation, was seen to be the key. This vision encompassed everything from multimedia personal computers (PCs) to the Internet to wireless (mobile phone) networks. At the same time, IBM's strategic focus shifted from selling products to selling a variety of services targeted at different industry groups, as well as selling a variety of information related services and using an extensive network of strategic alliances as a source of new targeted software products, which was expected to add \$10 billion in IBM sales by 2003 [Gerstner 2003; Lohr 2002(A)].

It was possible that in 2002, with the appointment of a new innovative "break the mold" CEO, Sam Palmisano, IBM would again strike out in new strategic directions, new directions which IBM will need to keep up the growth pace achieved under Gerstner [Ante 2003; Kirkpatrick 2002]. For example, in mid-2002 IBM

bought PricewaterhouseCoopers' consulting business for \$3.5 billion continuing its strategic move to strengthen its technology-services and consulting business and shift away from its computer-manufacturing roots [Bulkeley and Dunham 2002]. Its IBM Global Services SBU which includes both consultancy and IT support services in 2002 accounted for 45 percent of the company's \$81 billion revenue [Schlender 2003]. *Consultants News* listed IBM as the top ranking consultancy firm in 2001 [Skapinker and Foremski 2002].

At the same time IBM announced the opening of a \$2.5 billion specialized chip plant, the largest single investment IBM ever made, in part in order to remain a leader in the slumping global semiconductor business [Lohr 2002(B)]. In late 2002, it also announced it was moving strongly into "partnering" agreements with outside companies for software development and was creating "on-demand" delivery of computer related technologies and services [April and Lelii 2003; McDougall 2002], as did Hewlett Packard [Thibodeau 2003]. IBM has since 2001 also purchased a number of software development and services companies worth \$1.5 billion [Ante and Kerstetter 2002], all part of its effort to transform itself over the long term and create value in a new era of outsourced business processes [London 2003]. In order to better serve its software customers, in late 2003 IBM announced a major reorganization of its \$13.1bn software business, the second largest in the world vertically around industries [Foremski 2003(A&B)]. At the same time, IBM itself like many other U.S. companies was planning major outsourcing of segments of its own businesses. [Bulkeley 2004].

This experience shows how a strategic leader can encounter problems and change strategies when trying to anticipate and respond to the perceived long-term future, since the risks are great in breaking new ground in a major way. Nonetheless, it is often necessary to try to attempt to "create" the future [Feder 1999; Prahalad 1999]. Many firms are aggressively exploring "real option" scenarios involving how the Internet and other digital technologies will be changing traditional industry relationships, for example, by bypassing intermediaries and selling directly to customers. As another example, the speed and efficiency of telecommunications networks will be increased greatly by recent developments in Ethernet technology [Markoff 2000], creating both threats and opportunities. Many new industries are expected to be created based on these so-called "digital" and other technological forces [Downes and Mui 1998].

This experience also illustrates some of the ways many businesses are going about breaking with past molds dramatically in their search for future trends and ways to exploit them – another example of creative

destruction – in an effort to deal with a rapidly changing market environment. Only time will tell whether or not, or to what degree, such individual company efforts are successful.

**CONCLUSION: A NEVER ENDING PROCESS OF CONTINUING EVALUATION – THE
OPPORTUNITY SEARCH FOR POSSIBLE WINNING STRATEGIES**

Building on (and continuing) the structured analysis of the present market and of possible future market trends, and on an analysis of a specific company's strengths and weaknesses and possible strategic moves which might affect its performance in that marketplace, this paper describes winning strategies other specific companies have developed in response to the observed and anticipated driving forces in competitive markets. The studies are designed to stimulate further thinking about ways to brainstorm new ideas and write new scenarios for possible alternatives a company under study might want to explore. Since a manager in a specific company formulates strategies within a specific competitive market, the strategic management process in practice requires studying both the kind of company under study and the industry/competitive market segments involved, as well as their impact on winning.

Such a detailed structured situation analysis in practice can be a challenging task for several reasons, since it involves not only gathering and reformulating information on the current situation, but also studying future trends (especially those related to the digital and knowledge revolutions, globalization, privatization and deregulation, and the abundance of capital). The analysis probes the impact of these new forces on opportunities and keys to success now and in the future -- especially in regard to changes in the value chain and the speed of change, the relative position of competitors and likely competitor moves, and the company being planned for. Additional professional expertise can be nurtured through becoming familiar with other company experiences and the available strategic bag of tools and how they were developed, tools which are available to help stimulate ways to create a specific solution tailored to his/her own situation -- the perspective controlling idea formulation.

In addition to focusing on doing as well or better than the competition in all areas at the threshold or survival level, a firm formulates specific winning strategies in those areas most affecting profitability and success in which a company has or can acquire the competencies to distinguish itself from the competition in dramatically differentiated better ways. Winning is going beyond meeting threshold requirements and just trying to do everything well or marginally better than or different from the competition in today's environment. The search is not for differences for their own sake; the differences must produce improved results in the marketplace. Winning

strategies formulated by different companies result from taking steps to effectively respond to the many forces driving change in competitive markets in the **anticipated intermediate future**, and over the **longer-term more uncertain future**, as well as in the **short-term**.

REFERENCES

- Anders, George, "There Goes the Status Quo," *The Wall Street Journal*, Technology Section, May 19, 2003, pp. R1, R4,
- Ante, Stephen E., and Jim Kerstetter, "IBM: A Matter of Self-Interest," *Business Week*, December 16, 2002, p. 42.
- Ante, Stephen E., "The New IBM," *Business Week*, March 10 2003, pp. 80-88.
- April, Carolyn A., and Sonia R. Lelii, "On Demand On The Horizon," *VARBusiness*, October 27, 2003, pp. 30-36.
- Austen, Ian, "In Germany, Customers Scan As They Shop," *The New York Times*, May 12, 2003, p. C7.
- Beard, Allison, "E-tail of Woe," *Financial Times*, January 6/7, 2001(A), p. xxv.
- Beard, Allison, "Glum Signals on Mobile," *Financial Times* January 6/7, 2001(B), p. xxv.
- Beard, Allison, "Martha Stewart Losses Could Hit \$9m," *Financial Times*, August 12, 2003, p. 13.
- Booth-Thomas, Cathy, "The See-It-All Chip," *Time*, Special Bonus Section, September 22, 2003, pp. A7-A16.
- Boyle, Matthew, "Brand Killers," *Fortune*, August 11, 2003, pp. 89-100.
- Brick, Michael, "Indications of New Vision at Microsoft," *The New York Times*, February 15, 2001, pp. C1, C8.
- Brown, Shona L., and Kathleen M. Eisenhardt, *Competing on the Edge: Strategy as Structured Chaos*, Boston, MA: Harvard Business School Press, 1998.
- Buckman, Rebecca, "Microsoft Net Doubles As Sales Rise," *The Wall Street Journal*, October 18, 2002, pp. A3, A7.
- Bulkeley, William M., "IBM Documents Give Rare Look At Sensitive Plans on 'Offshoring'," *The Wall Street Journal*, January 19, 2004, pp. A1, A8.
- Bulkeley, William M., and Kemba Dunham, "IBM Speeds Move to Consulting With \$3.5 Billion Acquisition," *The Wall Street Journal*, July 31, 2002, pp. A1, A6.
- Cane, Alan, "'Fables' Phenomenon Heralds New Era," *Financial Times*, FT/IT Section, April 17, 2002, p. 11.
- Chang, Huang Chin, and Robert J. Mockler, *Taiwan Semiconductor Manufacturing Company (TSMC): The Semiconductor Industry (IC Foundry)*, New York: Strategic Management Research Group (St. John's University), 2000.
- Coffey, Brendan, "Publish or Perish," *Forbes*, November 13, 2000, pp. 252-256.

Cole, George, "News Spreads by Word of Mouth," *Financial Times*, April 12, 2000, p. 11.

Coleman, Calmetta, "Bank One to Fold WingspanBank.com Into Its Online Unit," *The Wall Street Journal*, June 29, 2001, p. A2.

Deutsch, Claudia H., and Barnaby J. Feder, "A Radio Chip in Every Consumer Product," *The New York Times*, February 25, 2003. pp. C1, C5.

Downes, Larry, and Chunka Mui, *Unleashing the Killer App: Digital Strategies for Market Dominance*, Boston, MA: Harvard Business School Press, 1998.

Editors, "Charles Schwab Posts 29% Rise n Profits," *Wall Street Journal*, July 24, 2003, p. C11.

Ewalt, David M., "Gillette Razors Get A New Edge: RFID Tags," *InformationWeek*, January 15, 2003(A), p. 22.

Ewalt, David M., "Wal-Mart Shelves RFID Experiment," *InformationWeek*, July 14-21, 2003(B), p. 29.

Feder, Barnaby J., "Plotting Corporate Futures," *The New York Times*, June 24, 1999, pp. C1, C25.

Foremski, Tom, "IBM To Shake Up Software Business" and "Big Blue Bets on a Coming Revolution," *Financial Times*, December 1, 2003(A&B), pp. 17 and 19.

Fuller, Thomas, "How Microsoft Warded Off Rival Software," *The New York Times*, May 15, 2003, pp. C1, C6.

Gasparino, Charles, and Ken Brown, "Schwab's Own Stock Suffers From Move Into Online Trading," *The Wall Street Journal*, June 19, 2001, pp. A1, A6.

George, Nicholas, "Second Quarter Curbs Halt Rise in Mobile Phone Sales," *Financial Times*, August 30, 2001, p. 1.

Gerstner, Louis V., Jr., *Who Says Elephants Can't Dance?*, New York: Harper Business, 2003.

Greene, Jay, "Microsoft: How It Became Stronger Than Ever," *Business Week*, June 4, 2001, pp. 74-85.

Hamel, Gary, and Jeff Sampler, "The E-Corporation," *Fortune*, December 7, 1998, pp. 81-92.

Hansell, Saul, and David D. Kirkpatrick, "Borders Is Said To Be Quitting Online Effort to Sell Books," *The New York Times*, April 11, 2001, pp. C1, C2.

Hansell, Saul, "United Airlines Buys Online Marketer," *The New York Times*, June 5, 2001, p. C4.

Johnson, Jo, and Robert Clow, "Vivendi to Accept Offers for Publishing," *Financial Times*, September 8, 2002, p. 8.

Kirkpatrick, David D., "Book Buyers Stay Busy But Forsake Bookstores," *The New York Times*, June 30, 2003, pp. C1, C7.

Kirkpatrick, David, "The Future of IBM," *Fortune*, February 18, 2002, pp. 59-70.

Konicki, Steve, "Sophisticated Supply," *InformationWeek*, December 10, 2001, pp. 22-24.

Landler, Mark, and Geraldine Fabrikant, "A Tortoise, Many Hares, and Web of Convergence," *The New York Times*, Money & Business, August 4, 2002, pp. 1, 11.

Lee, Louise, "Online Grocers: Finally Delivering the Lettuce," *Business Week*, April 28, 2003, p. 67.

Lohr, Steve, "He Loves to Win, At I.B.M., He Did," *The New York Times*, Money & Business Section, March 10, 2002(A), pp. 1, 11.

Lohr, Steve, "I.B.M. Opening A \$2.5 Billion Specialized Chip Plant," *The New York Times*, August 1, 2002(B), pp. C1, C4.

Lohr, Steve, "World-Class Chip, but a Fragile Business," *The New York Times*, August 4, 2003(A), pp.C1, C4.

Lohr, Steve, "You Call This a Midlife Crisis?," *The New York Times*, Money & Business, August 31, 2003(B), pp.1, 9.

London, Simon, "Radio ID Tags Spread Waves of Anger Among Privacy Activists," *Financial Times*, March 1/2, 2003, p. 22.

Lowry, Tom, et. al., "Mega Media Mergers: How Dangerous?," *Business Week*, February 23, 2004, pp. 34-42.

Maitland, Alison, "Giant Leap For Human Resources," *Financial Times*, March 27, 2002, p. 9.

Markoff, John, "Ethernet Finds a New Level," *The New York Times*, June 5, 2000, pp. C1, C3.

Markoff, John, "Hollywood Mogul Plays By Technology Rules," *The New York Times*, February 3, 2004, pp. C1, C10.

Marsh, Peter, "Make Every One Different," *Financial Times*, Manufacturing Excellence Section, May 21, 2001, p. III.

Martin, Peter, "The Shape of Things to Come: What Will Happen to Internet Companies," *Financial Times*, December 30, 2000, p. 9.

McDougall, "Real Time For Fast Times," *InformationWeek*, November 4, 2002, pp. 18-20.

Mockler, Robert J., *Multinational Strategic Management*, Birmingham, New York: The Haworth Press, 2002(A).

Morrison, Scott, "The Growth of the Shrinking Computer," *Financial Times*, July 16,2003, p. 9.

Mullaney, Timothy J., Heather Green, Michael Arndt, Robert D. Hof, and Linda Himelstein, "The E-Biz Surprise," *Business Week*, May 12, 2003, pp. 60-68.

Murphy, Chris, and Mary Hayes, "Tagline," *InformationWeek*, June 16, 2003,pp. 18-20.

Neuborne, Ellen, "Coaxing With Catalogs," *Business Week :E-biz*, August 6, 2001, p. EB6.

O'Connor, Ashling, "Reed Gives Away OAG As Net Overtakes Travel Bible," *Financial Times*, July 5, 2001, p. 15.

Pollack, Andrew, "The Genome Is Mapped. Now He Wants Profit," *The New York Times*, Money & Business Section, February 24, 2002, pp. 1, 14.

Porter, Michael, *Competitive Strategy*, New York: The Free Press, 1980.

Prahalad, C. K., "Changes in the Competitive Battlefield," *Financial Times*, Mastering Strategy Section, Part Two, October 4, 1999, pp. 2-4.

Richtel, Matt, "EToys to Lay Off Its Remaining Workers," *The New York Times*, February 6, 2001, p. C8.

Richtel, Matt, "MGM Mirage Seeks to Lure The Gamblers of Cyberspace," *The New York Times*, September 28, 2002, C1, C14.

Ridgway, Nicole, "Down to the Wire," *Forbes*, November 12, 2000, pp. 222-228.

Roberts, Dan, and Edward Luce, "Service Industries Go Global: Skilled White-Collar Jobs Are Starting To Migrate To Lower-Cost Centres Overseas," *Financial Times*, August 20, 2003, p. 11.

Romero, Simon, "Against the Tide On the Mobile Internet," *The New York Times*, January 1, 2001, p. C3.

Rose, Matthew, "Martha, Martha, Martha," *The Wall Street Journal*, June 19, 2002, pp. B1, B5.

Russell, Joy D., "Banking on Technology," *VARBusiness*, March 18, 2002, pp. 77-80.

Schlender, Brent, "How Big Blue Is Turning Geeks Into Gold," *Fortune*, June 9, 2003, pp. 133-140.

Schwartz, Jeffrey, "Emerging Technologies," *VARBusiness*, March 18, 2002, pp. 53-68.

Shirouzu, Norihiko, "Ford's New Development Plan: Stop Reinventing Its Wheels," *The Wall Street Journal*, April 16, 2003, pp. A1, A14.

Skapinker, Michael and Tom Foremski, "Bigger Blue," *Financial Times*, August 2, 2002, p. 10

Skapinker, Michael, and Victoria Griffith, "The Survival Expert," *Financial Times*, June 5, 2000, p. 8.

Smith, Craig S., "Across China, New Economy Folds Into Old," *The New York Times*, July 6, 2001, p. C1.

Tam, Pui-Wang, "Cloud Over Sun Microsystems: Plummeting Computer Prices," *The Wall Street Journal*, October 16, 2003, pp. A1, A16.

Tedeschi, Bob, "A Lands' End Experiment in Selling Custom-Made Pants Is a Success, Leaving Its Rivals to Play Catch-Up," *The New York Times*, September 30, 2002, p. C7.

Tedeschi, Bob, "Online Businesses Are Supplementing Their Fleet E-mailings With An Old Warhorse: Direct Mailings," *The New York Times*, August 6, 2001, p. C4.

Thibodeau, "HP's 'Adaptive' Strategy Faces Challenging Market," *Computerworld*, May 12, 2003, p. 4.

Towers, Basil, "Company Power in a Nutshell," *Financial Times*, July 30, 2001, p. 8.

Trachtenberg, Jeffrey A., "Borders Sets Out To Make the Book Business Businesslike," *The Wall Street Journal*, May 20, 2002, pp. B1, A6.

van Grinsven, Lucas, "Triple Alliance Seeks Fast Track For M-Commerce," *Financial Times*, April 12, 2000, p. 1.

Welch, David, with Kathleen Kerwin, "Detroit Tries It the Japanese Way," *Business Week*, January 26, 2004, pp. 75-77.

Westbrook, Roy, "Production: Look to the Process For a Better Product," *Financial Times*, Mastering Management: Part Three, October 16, 2000, pp. 6-7.

Wind, Jerry Yoram, and Jeremy Main, *Driving Change*, New York: The Free Press, 1998.

NEGOTIATION PARADIGMS IN RUSSIA : FIVE EMPIRICAL STUDIES IN BUSINESS NEGOTIATION

Guy A Deloffre
ICN Business School, Nancy, France
guy.deloffre@icn-groupe.fr

ABSTRACT

This paper presents the results of several years of teaching and consulting actions in Eastern Europe and particularly Russia: it brings its contribution to the subject through the results of empirical case studies, and gives an attempt to question existing models to describe specificities of Russian business life. Through five cases, several dimensions, such as time, information, power, or language, complexity of issues, phases and tactics are questioned to describe how Russians understand negotiation.

INTRODUCTION

Since the fall of the Berlin Wall and later the Russian Perestroïka in 1991, Western companies have intensely developed their operations towards Eastern Europe, not only towards the recent new members of the European Union, like Poland, the Czech or the Baltic Republics, but also towards countries like Russia, Ukraine, etc. The activities which have been developed include sourcing of raw material and products, local production of goods, distribution of western and eastern products, establishment of services, financial products, etc.

However experience shows that it is not easy for Westerners to adapt fully to the characteristics of the Russian system where traditional and modern aspects coexist more or less harmoniously. Following the recent development of Western companies towards east European countries, specific aspects of communication and negotiation have shown the limits of existing models to describe and explain them : how to understand the people and especially the managers of these countries, how to deal with their differences, how to successfully work together and negotiate.

This paper presents the results of several years of teaching and consulting actions in Eastern Europe and particularly Russia: it brings its contribution to the subject through the results of empirical case studies, and gives an attempt to question existing models to describe specificities of East European business life.

The first case describes what happens when two Russian companies work together and explores what their managers consider as appropriate in negotiation tactics.

The second case analyses how a private Russian company prepares and controls the negotiation when it has to deal with a public organisation (like a town or a region).

The third case brings a contribution to describing human resources management issues, such as recruiting, in Russian companies.

The fourth case brings reflexions about Russian industry today and some aspects of work which are not traditionally evident, like responsibility or internal contracting.

The fifth case describes what happens when a Russian company has to negotiate with its sub-contractors and suppliers: how asymmetry of power has an influence on negotiation. The conclusion of these studies tries to question existing models about negotiation, and is a link with some aspects of the Russian mentality as a means of explaining some of the permanent behaviours of the Russians.

FIRST SITUATION

The first situation happens in an industrial company where several teams have to cooperate and work together in order to solve technical problems; the negotiators are managers of teams, they discuss problems linked to production through internal negotiations: meetings with their subordinates and colleagues.

The observations: during technical meetings, I had the opportunity to observe three specific behaviours:

- The first one consists in negotiating the negotiation itself: one manager from one division proposes to discuss, to negotiate, and his counterpart, from another division, does not accept. He replies: "we are going to (work, to decide by our self), and after, we will speak (together) »;
- The second one is an exchange process. It happens during a meeting, later, where the managers exchange technical positions and technical positions: "if we do this (solution), the you should do that (solution) and this will be acceptable»;
- The third one is related with preparing the meeting. One of the managers comes to the discussion table with a pre-prepared solution which he presents as the only valid one: "This is what you have to do". He does not make any difference between an opening position and a final position : the opening position is not seen as a base for mutual construction through discussion, but more as the result of a previous calculation which has to be accepted as such by the other side : "we calculated, and this is the solution you will have to accept".

SECOND SITUATION

The second situation is based on comments made by some managers working in an international industrial company who have to negotiate with a public organisation (equivalent to a town or an urban community). The subjects of the talks involve technical, industrial aspects, like the development of a plant, the strategy of the group, etc. as well as social or political aspects: employment, economic development.

When analysing how they negotiate, these executives describe how they prepare the talks and how they keep in mind specific interests.

They declare: "We always negotiate in teams of 2, sometimes more than 2 persons"... "Previously to the talks, we conduct a study about the persons who will be in front of us ; we visit internet sites, we try to find the curriculae..." ; what is their age, their exact position, their academic background, who is a self made man, who studied business, etc.

"The one who gathers his own interests and the interests of his boss, this one is tough to deal with... ; ...we always take great care of the social interests, of the political goals of the town...".

"During the negotiation, we sign protocols - not a contract, but what we agree upon..."
Political interests are not seen the same way from a Western country and from Russia. "Since we are part of a Western group, the (relative) importances, internal and external, are not the same..."

THIRD SITUATION

The third situation takes place in several companies of different sizes, big and small; I present here the result of several interviews with managers in charge of human resources management and the cohort of their associated events: recruiting, training, etc.

When describing the recruitment policy of a company, one manager explains that the salary should not be the only element to negotiate: the salary itself is formed of basic wages and premiums, bonuses, so all aspects should be negotiated within a global financial approach. Moreover, non financial elements should also be considered, such as lodging, training programs, or health insurance or even funding for pensions...

In a big company, subsidiary of a Western group, a manager in charge of recruiting explains the politics of her company: "We know that we do not have very high salaries, but the company itself represents a sort of warranty; about the salary, we propose the salary and he accepts or not...»; she explains that she has no margin to negotiate the salary when recruiting new personal.

When considering salaries in Russian industry, the average salary lies around 300 Euros monthly (500 USD) in 2007 in Russian companies. It lies around 500 Euros (700 USD) in Russian branches of Western groups and the general tendency is to rise slowly but regularly. One of the main difficulties for companies is not only to find competencies (competent people, like people who were trained in Western countries are scarce resource), but also to retain them: because of low wages and the general lack of effective human resources management policies, the turnover of workers is important.

FOURTH SITUATION

The fourth situation has been observed inside a company where several departments must work together. Internal supplier-client relationships are common in business today, but they imply some aspects of negotiation which are not always understood or accepted by the negotiators: continuous amelioration means that many people must be implied in the process, and particularly that the manager is ready and willing to listen to his subordinates. He has to accept the principle that subordinates can have good ideas.

It is necessary to limit the number of hierarchy levels in order to keep the organisation as simple as possible, so that the problems are solved where they arise, instead of being pushed upwards to the upper levels. It is also necessary to bring colleagues to cooperate together in order to solve complex problems and situations, even if no hierarchy link means that no one can force the other one to cooperate.

An other paradox consists in balancing between a perfect standardisation of tasks, methods and

procedures, and a maximum of flexibility in communication and negotiation. On the one hand, industrial quality means a maximum of control over all methods; on the other side, communication must be active and visible (through public posters, charts, etc.) ; subordinates must be given a possibility, either to interfere with processes (like the possibility of stopping production in the case of a problem affecting all products), or to contribute to them, through proposals. It is then important to show them that their proposals are interesting and useful.

FIFTH SITUATION

The fifth situation develops within a cascade of technical negotiations between an industrial company and its first tier suppliers. The company makes technical products and uses for the assembly process parts which are sourced from external suppliers. Since there is a quality insurance process at the client, representatives (they are called here consultants) of this company have to visit the suppliers in order to evaluate the quality level of the processes, the products and their production, and in order to check if the organisation at the supplier's premises is conform to the quality norms published by the client and imposed to all its suppliers.

This leads to 2 types of negotiations: a first one, where the consultants negotiate with the representatives of the supplier, and a second one, where the manager of the consultants has to evaluate their job from outside.

In the first negotiation, one can see that the consultants try, not only to represent, but also to impose the views of the client. Therefore they tend to adopt a directive and rather one-sided approach towards their counterparts. If this can be justified for technical reasons because the application of norms demands for it, it often leads to social difficulties and conflicts at the supplier: the brutal intrusion of the consultants into their activity is viewed as unacceptable and even sometimes intolerable. The result of these frictions is partly mis-analysed by the consultants, who declare: "They don't like us...", not establishing a clear difference between the technical negotiation and the social communication.

Back home, they must negotiate with their manager: report their action and prepare next visits. While doing this, two aspects emerge: the first one is the technical aspect of the supplier-client relation, based on quality control, quantities produced, procedures, deliveries, delays, etc. The second one, more difficult to apprehend, is the social aspect: not only the difficult personal relationships with the representatives of the supplier, but also the symbolic and psychological aspect of two organisations that have to work together. In that case, the "They don't like us" becomes "The client is impossible". And one of the tasks of the manager consists then in negotiating with low moral subordinates to motivate his troops.

DISCUSSION

Through these practical examples, it is plausible to draw global conclusions about negotiation in Russia. Beyond widely used models, such as Hall (59), Hofstede (91) or Trompenaars (93), I used two specific models to do this. These models do not mainly focus on cultural dimensions, but more on technical elements which structure negotiation.

The first model is based on a trilogy of concepts: time, information, power. It focuses on the

process of negotiation, where time, information and power are resources used to bring the negotiation to go along, or to gain tactical advantage over the counterpart.

Time

Russian time in negotiation is linked with acceleration and not only with duration. Russian people are said to be polychronic negotiators (Hofstede, 91), but what is characteristic of the cases described is the acceleration of time before and during the negotiation: preparation is an active phase, it is slow and deep: all the possibilities, all movements are explored, all the arguments prepared. Negotiating, on the contrary, is a rapid phase, more active, where all the cognitive material is used rapidly in order to obtain a result.

Information

Information is considered a powerful tool in Russia. It is the base of any negotiation, and there is ample evidence of this through the careful preparation presented in situation nr 2. Our experience led to interesting constataions: some Russian negotiators have a tendency to ask all what they want to obtain, they put all the informations and details about their position "on the table" at the beginning of the talks.

Power

It is not common for Russian negotiators to accept the idea that a negotiation can be equal, fair: they present a permanent tendency to consider that power in a negotiation can only be unequally distributed. Therefore, they think that there must be a strong and a weak partner (which does not necessarily mean a winner and a loser). A permanent concern is therefore: how to gain advantage over the counterpart, through power, or authority, or influence. On the other side, Russian negotiators have sometimes a tendency to consider themselves as "victims" when they are not in a powerful position, especially in front of Western negotiators.

The second model presents some elements common to many negotiation situations and specific to Russian approach of talks. It is part of a balanced approach (Czeslaw, 2005). Some of its constituents are listed below.

Language

How to understand the concept of negotiation ? If we compare French or Spanish acceptions of the word with the Russian one, then we can observe that :

- The French "négociation" uses the core "négoce", which means commerce, trade ; this approach is comparable to the Spanish, where "negociacion" is linked to "negocios". These words mean exchange, commerce, trade, and include an implicit dimension of business, of professional activity ;
- The Russian "peregovory", which stands for negotiation, uses the core "govor" which means speak, talk. In that case, the implicit meaning is not the one of trade, commerce, but more of talks, discussion, a social activity. The exact equivalent of the Russian "peregovory" should be the French "pourparlers", which means talks and is used mainly for political and diplomatic situations.

This is related to the fact that many Russians consider negotiation more as a social activity (speak means speak together, speak with the others) rather than a technical one.

Complexity of issues

The situations listed and described here involve many issues in the negotiations (Raiffa, 82 : one issue - many issues). Internal and external issues are the most complex elements to deal with during talks, opposition between them often makes decision difficult to reach.

Besides, social issues are clearly part of any negotiation process, even if this one is supposed to be only professional : it is impossible to reach an important agreement with Russians without developing a personal relationship with individuals. In Russia, social factors are often preeminent over technical issues.

Use of phases and tactics

Russian negotiators emphasize preparation before any negotiation ; they dig in all the facts, prepare scenarios, evaluate the relative positions, try to gain some advance from the context. Opening phases are crucial in order to take a tactical advantage ; however, discussion and exchange phases are less constructive : they do not consist in building a common project, but instead, in playing and reorganising positions against each other.

When discussing their tactics, for example during interviews, Russian negotiators describe their approach in terms of a prescriptive one, more than a descriptive one : how to negotiate, how to do it better, how to bring the other side to a clear issue. Russians are also excellent chess players...

Limits and perspectives

We limited our study to industrial situations, to small and big private companies, to negotiations where the majority of players are men (even if some of our contacts were women). When discussing with managers trained in Western countries, we did not take into consideration the country where they studied.

Existing models about cultural behaviours and cultural differences are not very useful in Russia. They are pertinent in Western countries because they have been thought and developed by Westerners for Westerners. But when they are used to apprehend situations where context, players or value systems are based on different principles, they bring few help in understanding and foreseeing the behaviour of the negotiators. This is why many Western businesspeople show a tendency to evaluate the behaviour of Russians as difficult to understand or even irrational.

CONCLUSION

It is possible to conclude this paper with a paradox: on one side, the "Russian" does not exist. The size of the country, the great diversity in local cultures, languages, religions, the influence of geography or climate, historical or political factors do not allow to apprehend the Russian as a unique character who would show homogenous negotiating behaviour and style.

But on another side, there are enough common aspects in this study to speak of a "Russian" in negotiation. Of course, this must be seen as a linguistic commodity within the framework of this study, but it opens also new development for further studies : it could be interesting to evaluate where the limits are, between global characteristics common to all Russians, and the aspects specific of some Russians only.

(References available upon request from Guy A Deloffre)

THE IMPACT OF HAVING BOTH: FRANCHISED STORES AND COMPANY-OWNED STORES WITHIN A SAME NETWORK ON THE INTERNATIONALIZATION: EMPIRICAL RESULTS FROM THE US AND FRENCH FRANCHISING NETWORKS

Rozenn PERRIGOT

IGR-IAE at the University of Rennes 1, ESC Rennes School of Business
Center for Research in Economics and Management (CREM UMR CNRS 6211)
11 r. Jean Macé, CS 70803, 35708 Rennes Cedex 7, France, rozenn.perrigot@univ-rennes1.fr

ABSTRACT

Two topics often explored – but separately – in the franchising literature are network internationalization and plural form. The purpose of this paper is to explore if there is any significant, positive or negative, link between plural form and network internationalization. The empirical study involving 898 US and French networks among which 42.1% are international highlights the existence of significant differences between international networks and purely domestic networks in terms of plural form. Indeed, the plural form rate for international networks is lower than this for domestic networks. Moreover, the results of the logistic regression models underline the significant and negative impact of the plural form on internationalization.

Keywords: Franchising, plural form, internationalization, multi-country study, USA, France.

INTRODUCTION AND RESEARCH PURPOSE

Two important topics often explored in the franchising literature are plural form and internationalization. On the one hand, conceptual papers (Bradach and Eccles, 1989; Dant *et al.*, 1992) and empirical papers (Lafontaine and Kaufmann, 1994; Bradach, 1997; 1998; Lafontaine and Shaw, 1999; 2005; Dant and Kaufmann, 2003) have contributed to the better understanding of the plural form, i.e. the coexistence of franchised units and company-owned units within a same network, and have often underlined its advantages in terms of costs, growth, quality, risk management (Ehrmann and Spranger, 2004), and so on.

On the other hand, according to Elango (2007), the literature about international franchising can be classified into three main streams. The first stream (Welch, 1989; Alon and McKee, 1999; Hoffman and Preble, 2001; Welsh *et al.*, 2006) focuses on macro differences across countries to determine countries in which franchising is spreading and gaining acceptance as an organizational form. The second stream of literature (Fladmoe-Lindquist and Jacque, 1995; Contractor and Kundu, 1998a; 1998b; Burton *et al.*, 2000; Erramilli *et al.*, 2002) focuses on the choice of entry mode when seeking international markets. The third stream (Walker and Etzel, 1973; Hackett, 1976; Aydin and Kacker, 1990; Huszagh *et al.*, 1992; Kedia *et al.*, 1994; 1995; Julian and Castrogiovanni, 1995; Shane, 1996) focuses on factors driving franchisors to international markets as well as factors distinguishing franchisors seeking to operate internationally from those who only focus on the domestic market.

Nevertheless, these two research subfields, i.e. plural form and franchising network internationalization, have not converged so far. The purpose of this paper is therefore to explore if there is any significant link between the plural form and the internationalization of the franchising networks.

One of the most influent papers about plural form is perhaps this of Bradach (1997; 1998). This author focused on plural form through an in-depth exploratory study of five US fast food networks. He explained that the plural form within a franchising network is aimed at meeting the following managerial challenges: spatial expansion with the addition of new units, brand protection by maintaining the concept uniformity, local reaction to threats or opportunities, and service and/or product concept evolution with the need to adapt the concept to changes. This researcher described several processes emerging in a plural form network and helping the franchisor to take on more easily the four challenges. They deal, for instance, with the additive process during the network development: the franchisor shows his/her know-how through the company-owned units, and thus attracts new franchisees, and also a franchisee can create his/her own units. A socialization process exists as well because the franchisor personnel can become a source of new franchisees. Moreover, there is a mutual learning process in a plural form network that facilitates the generation, test, selection and implementation of new ideas. Following this pioneer research, several researchers have highlighted the advantages of the plural form. Ehrmann and Spranger (2004) classified these research papers according to the different motivations for the franchisor to apply plural form. They deal with cost aspects, growth aspects, quality aspects and risk aspects.

The literature review and the synergies provided by mixing franchising and company arrangement within a same network contribute to formulate the two following propositions: *(P1) There is a significant influence of plural form on network internationalization*, and *(P2) The influence of plural form on network internationalization is positive*.

These two propositions are tested using a multi-country approach, and more specifically choosing the US and French franchising industries as the context because of their dynamism. In the USA, there are 1,500 networks including 767,483 franchised and company-owned units and generating US\$ 624.60 billion. Furthermore, France has been selected because of its paramount situation at the European level as far as the number of networks is concerned. There are 929 networks in the French territory, with a total of 69,339 franchised and company-owned units generating € 134.7 billion (French Federation of Franchising, 2007).

It can be noticed that most of the previous research on plural form have been single-country research (Lafontaine and Shaw, 1999; 2005; Cliquet, 2000; Frazer, 2001; Lopez and Gonzales-Busto, 2001; Dant and Kaufmann, 2003; Ehrmann and Spranger, 2004; Windsperger, 2004; Windsperger and Dant, 2006), except a recent exploratory analysis comparing French and Brazilian franchising case studies (Azevedo and Silva, 2005), and a forthcoming paper about a cross-cultural comparison of the plural forms in franchise networks with a focus on the USA, France, and Brazil (Dant *et al.*, forthcoming), both papers underlining the need for future data-driven cross-cultural investigations of the plural form. It is attempted in the present paper to contribute to fill this gap with exploring the plural form and the network internationalization in the USA and in France.

The sample includes 405 US networks among which 239 networks (59%) are international ones, and 493 French networks among which 139 networks (28.2%) are international ones.

So, globally, combining the US and French data, 378 networks (42.1%) are internationalized whereas 520 networks (57.9%) are purely domestic networks.

As far as the methodology is concerned, descriptive statistics -and specifically t-test and Levene statistics- first highlight the significant differences that exist between purely domestic networks and international ones as far as the plural form is concerned. Other differences are also studied in terms of network age, network size, financial conditions of the franchising contract, etc. Secondly, logistic regression models enable analyzing the influence of plural form on network internationalization. Additionally, this multi-country study enables to highlight the main differences between the US networks and the French ones in terms of internationalization and plural form.

The paper is organized as follows. In the next section, the research design is presented. The results of the t-tests, Levene statistics and logistic regression models are then respectively displayed and discussed in the following sections.

RESEARCH DESIGN

Sample

The two research propositions are tested within a multi-country perspective. Indeed, the empirical study deals with the US and French franchising markets. These two countries have been selected because of their dynamism in terms of franchising activities. On the one hand, in the USA, there are 1,500 networks including 767,483 franchised and company-owned units and generating US\$ 624.60 billion. On the other hand, France is at the paramount situation at the European level as far as the number of networks is concerned. There are 929 networks in the French territory, with a total of 69,339 franchised and company-owned units generating € 134.7 billion (French Federation of Franchising, 2007). The source for the US data is the “*2006 Entrepreneur Magazine - Annual Franchise 500*”. This for the French market is the franchising directory entitled “*2006 Toute la Franchise, les Textes, les Chiffres, les Réseaux*” edited by the French Federation of Franchising. A total of 898 networks compose the final sample: 405 are US networks and 493 are French ones.

Methods

In a first step, descriptive statistics, and specifically t-test and Levene statistics, are used to highlight the significant differences – mainly in terms of plural form – that exist between purely domestic networks and international ones in the two markets, and also in the global database. In a second step, logistic regression models are run to analyze the factors for network internationalization. Here as well, the focus is mainly made on the influence of the plural form rate on network internationalization in the USA and France, and also globally.

Variables

The dependent variable in the logistic regression models is internationalization. It is a categorical variable coded 1 if the network is international and 0 if the network is purely domestic. In the US sample, a total of 239 networks (59%) are international networks, the 166 other networks (41%) are purely domestic ones. In the French sample, a total of 139 networks (28.2%) are international networks, the 354 other networks (71.8%) are purely

domestic ones. And globally, combining the US and French data, 378 networks (42.1%) are internationalized whereas 520 networks (57.9%) are purely domestic networks. The plural form rate (PFR) is the independent variable of the logistic regression models that enables to test the research propositions. The plural form rate corresponds to the percentage of company-owned units within the network. The average plural form rate is equal to 6.63% with a standard deviation of 15.43 in the population of US networks (N = 404), to 40.83% with a standard deviation of 35.00 in the population of French networks (N = 493) and to 25.43% with a standard deviation of 32.70 in the population of US and French franchising networks all together (N = 897). Based on past research findings, the following variables are used as control variables: network age in years, total network size in the domestic market, franchising fee in thousand dollars, royalties in percentage of the total sales, contract duration in years and the industry type.

MAIN RESULTS

All the tables of the results presented in this section are available upon request from Rozenn PERRIGOT (rozenn.perrigot@univ-rennes1.fr).

Results of the t-tests and the Levene statistics

According to the results of the t-tests and the Levene statistics, (1) at the global level, i.e. US networks and French ones composing a same database, international networks significantly have a smaller plural form rate (16.82%) than purely domestic networks (31.66%); (2) at the French level, international networks significantly have a smaller plural form rate (34.59%) than purely domestic networks (43.28%); and (3) at the US level, the international and purely domestic networks respectively have a plural form rate of 6.44% and 6.90%, but the difference is not significant. Other significant differences are found: international networks are older and larger than purely domestic networks in the USA, in France and at the global level. Nevertheless, there is not any significant difference between international and purely domestic networks as far as franchising fee, royalties and contract duration are concerned, except for the contract duration at the global level (9.35 years for the international networks and 7.57 years for the purely domestic networks).

Results of the logistic regressions

Models 1, 3 and 5 included only the control variables respectively for the US market, the French one and at the global level, whereas models 2, 4 and 6 allows testing the propositions about the significant and positive influence of plural form on internationalization respectively for the USA, for France and at the global level. The pseudo R² of Model 2, 4 and 6 whose values are respectively equal to 18.2%, 34.9% and 28.9% are satisfactory. Moreover, 66.9%, 80.6% and 72.9% of the networks are considered well classified respectively with models 2, 4 and 6. The variable PFR has a significant impact on internationalization at the 0.01 level for the French sample of networks and at the global level but not for the US sample of networks. P1 then receives statistical support in the French territory and at the global level but not in the US territory. Furthermore, this impact is negative. Indeed, the plural form rate is inversely related to the decision to internationalize the network in the three cases. P2 is therefore contradicted in the case of the three samples.

As far as the control variables are concerned, the variables AGE and SIZ have a significant and positive impact on internationalization at the 0.01 level in the USA, in France and at the global level. Older and larger networks are more likely to be international than new and small networks in the three samples. Nevertheless, FEE, ROY and DUR do not have any significant impact on internationalization, such as the industry type.

DISCUSSION

Discussion of the findings

Preliminary comments emerging from this comparative study of the US and French franchising industries deal with the differences observed in terms of the degree of internationalization of the franchising networks and the plural form rate. First, US networks are much more internationalized than French ones. Indeed, 59% of the US networks are international ones whereas only 28.2% of the French networks are international ones. Second, US networks are much less plural than French ones. The average plural form rate is equal to 6.63% in the population of US networks whereas it is equal to 40.83% in the population of French networks.

Furthermore, the empirical study involving 898 US and French networks among which 42.1% are international networks highlights the existence of significant differences between international networks and purely domestic networks in terms of plural form. First, the average plural form rate for the combined sample of US and French networks and having operations abroad is equal to 16.82% whereas this for purely domestic networks is equal to 31.66%. Second, in France, international networks significantly have a smaller plural form rate (34.59%) than purely domestic networks (43.28%). Moreover, the results of the logistic regression models for the total sample of franchising networks and also in the case of the French networks underline the significant and negative impact of the plural form on internationalization. *P1 (There is a significant influence of plural form on network internationalization)* is therefore supported whereas *P2 (The influence of plural form on network internationalization is positive)* is contradicted at the global level and in the French market. Third, in the US market, the international and purely domestic networks respectively have a plural form rate of 6.44% and 6.90%, and the difference is not significant. The influence of the plural form rate is negative but insignificant in the logistic regression model in the case of the US networks.

The findings of the empirical study dealing with the French market, and those at the global level, can seem surprising because the previous literature on plural form has highlighted the synergies (Bradach, 1997; 1998; Ehrmann and Spranger, 2004) provided by mixing franchising and company ownership within a same network. One could have thought that these synergies could have lead to greater probabilities to internationalize the network. On the one hand, one could have thought that a strong position in the domestic market, through many company-owned units, would have favored the network internationalization because of knowledge acquisition, etc. On the other hand, one could have thought that plural form would have facilitated the choice of entry modes to enter foreign markets. At the opposite, the empirical study underlines that the more the network has company-owned units, the less it is international. It seems that there is a substitution effect. In other words, if the franchisor invests in the network through the ownership of units, he/she will not explore the opportunities offered by foreign markets. This substitution effect deals with financial

resources, human resources, and efforts in terms of stimulation, control, etc. Furthermore, the risk-taking orientation can contribute to explain the simultaneous preference for franchising and internationalization.

Managerial implications and research contributions

As far as the managerial implications are concerned, this paper underlines the fact that when the franchisor is focused on company ownership in his/her domestic market, he/she is less likely to internationalize his/her network. Yet, this strong position in the domestic market through the company-owned units that provide information and knowledge about the local environment, that enable to test the innovations and train the new franchisees, etc. could favor the internationalization process. Indeed, the franchisor could use this experience acquired through company ownership to go abroad, either directly for instance asking some managers of company-owned units that already know the network principles to manage units in the new countries, or through franchising using this strong position in the domestic market to attract and recruit new foreign franchisees. This strong position in the domestic market could serve as a signal to build a brand image in foreign markets. A substitution effect has been previously mentioned. Nevertheless, the franchisors would not have to consider company ownership and internationalization as two opposite strategic directions. There is a learning effect: the knowledge directly acquired in the domestic market through the managers of the company-owned units can serve to expand the network abroad.

Moreover, the franchisors when entering a new market have to consider the differences in terms of network management that exist between them and the local competitors. For instance, US franchisors who enter the French market have to face French competitors that are more focused on the local market and also more oriented towards company ownership.

This paper contributes to enlarge the literature on plural form in two ways. First, contrary to many papers that have underlined the advantages of the plural form, this paper highlights a kind of limitation of the plural form in the sense that more the network is plural and less it is internationalized. The findings of the present research are consistent with some figures provided by Dant *et al.* (forthcoming) in a recent paper about a cross-cultural comparison of the plural forms in franchise networks. These authors noticed that the plural form rate of international networks was lower than this of the purely domestic networks. More specifically, working on 1,318 networks from the USA, France and Brazil, they found that the average plural form rate was equal to 31.24% for the purely domestic networks and to 20.03% for the international networks. The same tendency was also found in these three countries taken separately. More research are needed to explore this surprising relationship. Second, as underlined by Dant *et al.* (forthcoming), multi-country comparisons are very useful for helping the franchisors in the network management and internationalization.

REFERENCES

The references are available upon request from Rozenn PERRIGOT (rozenn.perrigot@univ-rennes1.fr).

AN INVESTIGATION OF CULTURAL WORK-RELATED VALUES IN SOUTH KOREA

Jennifer Ann Swanson, Dept. of Business Admin., Stonehill College, 320 Washington St.,
Easton, MA 02357; (508) 565-1349; jswanson@stonehill.edu

Kellyann Berube Kowalski, Dept. of Management & Marketing, Charlton College of Business,
Univ. of Massachusetts Dartmouth, 285 Old Westport Road, N. Dartmouth, MA 02747;
(508) 999-8327; kkowalski@umassd.edu

ABSTRACT

This working paper focuses on cultural work-related values in the Republic of Korea (South Korea) and their potential affect on multinational operations. More specifically, we discuss cultural values and gender roles in an attempt to increase the success rate of US expatriates working in South Korea and other professionals engaged in international business dealings in South Korea.

INTRODUCTION

The rapidly growing global marketplace has created many opportunities for American companies. Furthermore, globalization has become an important factor in an organization's success. As more and more companies create strategic alliances and enter into joint ventures in other countries, the frequency of international business transactions and the number of expatriate assignments are steadily increasing (Roy & Dugal, 1999). In particular, the number of expatriates being sent to Asian countries, including South Korea has increased significantly over the past ten years.

Since the early 1960s, South Korea has achieved remarkable economic growth. Although South Korea experienced the Asian financial crisis in the late 1990s that exposed certain economic weaknesses and recently had a downturn in consumer spending, they are considered a very solid economy with rapid export growth, moderate inflation, low unemployment, an export surplus, and a fairly equal distribution of income (U.S. Government, 2007). South Korea is a key player in the world economy and is considered one of the world's most technologically and scientifically advanced countries. It is a global leader in electronics and technological products and is successful in car manufacturing, shipbuilding, steel production, robotics, and biotechnology industries (U.S. Government, 2007).

The South Korean government opened up their markets after the Asian Financial crisis, ending the dominance of chaebols in the economy. This economic liberalization as well as an incentive program for foreign investors outlined in the Foreign Investment and Promotion Act of 1998 helped spur the growth of foreign direct investment inflows (fDiMagazine.com, 2007). Investments have been primarily in the manufacturing and service sectors with electronics, chemicals and automotives being the most attractive to foreign investors in the manufacturing

sector and insurance, financial services, retail and logistics leading in the service sector (fDiMagazine.com, 2007).

Currently, the United States is the largest foreign investor in South Korea. In addition, talks are underway for creating a free trade agreement between Korea and the US since they are such important trading partners (fDimagazine.com, 2007). Therefore, the ties between the two countries should continue to be strong in the future. The frequency of international business transactions and the number of expatriate workers in South Korea has been steadily increasing and should do so well in the future

This study on South Korean values will provide practical information for expatriates (and their families) and those conducting business with South Korea. In addition, it will provide updated information and insight into a fascinating country.

EXPATRIATES

Unfortunately, many of the international business dealings and expatriate assignments in South Korea may not be entirely successful as many expatriate assignments fail. One study uncovered that anywhere between 16 and 40 percent of all American employees sent abroad to developed nations return from their assignments early and almost 70 percent of employees sent to developing nations return home early (Shay & Bruce, 1997).

The International Orientation Resources, a HRM consulting firm, reported that 60 percent of expatriate failures occur due to one of three reasons – the inability of the manager to adjust to the new culture, the inability of the spouse to adjust to the new culture, and other family problems (Solomon, 1994). Therefore, successful expatriate placement depends on the expatriate worker's ability to understand and adjust to the new cultural values and norms of his/her host country (Hogan & Goodson, 1990; Flynn, 1995). Therefore, knowledge of the host country, including its culture, is crucial since expatriate workers must be able to adapt and assimilate to their new environment if they are to be effective in the workplace. Furthermore, the family of the expatriate worker must be able to adapt as well. Studying the host country prior to the assignment can help in the assimilation process. Pre-departure training for the expatriate and his/her family can help lead to cross-cultural awareness, sensitivity and adaptability. Cross-cultural training and country orientation programs have been cited as critical expatriate success factors, thus improving the adjustment process (Hogan & Goodson, 1990; Katz & Seifer, 1996).

CULTURE

According to Geert Hofstede, culture can be defined as “the collective programming of the mind that distinguishes the members of one group or category of people (i.e., nation) from another” (1997: 5). Culture affects the way we think, feel, behave, perceive the world, and behave. Culture helps to determine individual value systems. A value is defined as an individual attribute and is a “broad tendency to prefer certain states of affairs over others” (Hofstede, 1980: 19).

Hofstede's landmark study focused on a culture's affect on values in the workplace. Hofstede (1980) found that nations differed on five main work-related value dimensions: (1) Power

Distance (PDI), (2) Individualism-Collectivism (IDV), (3) Masculinity-Femininity (MAS), (4) Uncertainty Avoidance (UAI), and (5) Long-Term Orientation (LTO) (Hofstede, 1993). Please refer to Hofstede's work for an overview.

Hofstede initially analyzed the values of employees in forty different countries across these five dimensions. The position of a country on these dimensions helps us to understand how the society operates, including its management processes in the workplace. These cultural values lead to certain work-related behaviors that have a strong influence on how a company operates, its success, and how it deals with individuals from other cultures.

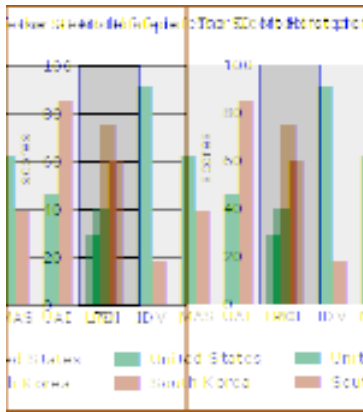
In analyzing South Korea across the value dimensions, Hofstede reported that it has a relatively high power distance (inequality was accepted), is moderately feminine, is highly collectivist, has a low tolerance for uncertainty and risk, and has a long term orientation (Itim International, 2007). The most important values in South Korea pertain to the Uncertainty Avoidance (UAI) dimension and the Individualist-Collectivist (IDV) dimension. This indicates that many South Koreans do not tolerate uncertainty, risk, or ambiguity. To reduce this uncertainty, strict rules, laws, policies, and regulations are adopted in an attempt to control outcomes. Change would not be readily accepted (Hofstede, 1980). Another important value was collectivism which indicates that there are close long-term commitments and relationships between individuals. Although this research will look at culture overall, this initial study focuses on the masculine-feminine dimension the most. Hofstede's results indicate that the Korean culture is somewhat feminine (Hofstede, 1980). However, the score does not indicate a high preference. Therefore, it will be interesting to see how South Koreans score on this dimension given the changes that have taken place in their culture.

In contrast to South Korea, Hofstede found that the United States has a low power distance, is masculine, is highly individualist, has a high tolerance for uncertainty and risk, and has a short term orientation (Itim International, 2007). Most Americans have very individualistic attitudes and the bonds between people are quite loose. Americans tend to be more self-reliant in the workplace and expect equality between the ranks. Since uncertainty is tolerated, there are fewer organizational rules need to minimize risk (Hofstede, 1980).

The cultural distance or intercultural disparity index (IDI) between South Korea and the United States is quite high. The IDI is based on Hofstede's cultural dimensions and is used to measure the relative distance of two given cultures and therefore the existence of relevant cultural differences. It is the extent to which the culture of the originating region (United States) differs from that of the host region (South Korea). As you can see from the chart on the following page, the IDI between the two countries is quite high.

High IDI translates into potential cultural clashes in the workplace since different cultural systems translates into different methods of communicating, different work habits and goals, different decision-making styles, and different motivational methods. As such, leadership behaviors have to be adjusted and adapted to the cultural value systems of the host countries. This further highlights the importance and need for cross-cultural training for those engaged in doing business in South Korea. The final paper will identify appropriate training modules as

well which will be useful for practicing managers in the development of organizational strategies and overall training programs.



*Chart: Comparison of United States to South Korea [US home country; S. Korea host country]
Source: Itim International, Geert Hofstede's Cultural Dimensions, 2007*

Although Hofstede's work is highly regarded and widely accepted, a few problems exist. One of the main problems is how dated the information is as data was collected in the late 1960s and early 1970s. As such, it is pertinent to update the cultural research on South Korea as this society has undergone enormous changes since Hofstede's study was conducted. Although this study will not be an exact replication of Hofstede's work, it will explore South Korean values in general and then focus on the masculine-feminine dimension. As mentioned above, South Korea is considered a culture that values feminine qualities, scoring a 39 on the masculinity dimension (on a scale from 0 to 100) with higher scores indicating a more masculine culture. In addition, we will investigate the values of the need for achievement, need for affiliation, and need for power as they are believed to be highly correlated with masculinity and femininity (Roy, 1999).

MASCULINITY AND FEMININITY

An individual's masculinity or femininity (also termed gender-role orientation) has an effect on how he or she will behave in an organization. Bem (1974) developed the Bem Sex-Role Inventory (BSRI) to measure individual gender-role orientations, which treats femininity and masculinity as completely separate dimensions allowing an individual to be high or low on both values. This is much different from Hofstede's survey, which treated masculinity and femininity as bipolar opposites. The characteristics listed for masculinity and femininity by Bem (1974) are similar to those used by Hofstede (1980) in his research.

Gender-role orientation can affect an individual's flexibility, leadership abilities, group behavior, gender stereotypes, and occupational choice. As such, an individual's gender-role orientation can help determine how he or she will behave in the workplace. Given the importance of gender-role orientation in the workplace, it is important to investigate this dimension in South Korea. Since Hofstede's culture study looked at masculinity and femininity as opposites, each country was

determined to be one or the other. Our study incorporates Bem's work and therefore the country and individuals within it can be classified as masculine, feminine, androgynous, or undifferentiated.

Based on the findings from Hofstede's work and the work related to the gender-role orientation, one might expect that the majority of individuals in South Korea are feminine in their value structure on the Bem Scale. However, this may not be the case given the dramatic changes over the past few decades in South Korea. As such, it is important to measure this characteristic.

NEED FOR ACHIEVEMENT, AFFILIATION, AND POWER

Two related values to masculinity and femininity are the need for achievement and the need for affiliation. Individuals who are masculine tend to have a high need for achievement whereas those individuals who are feminine are more likely to have a high need for affiliation (Roy, 1999; Hofstede, 1980). The same may hold true for the country level of analysis. Countries that value feminine traits tend to have a higher percentage of their population reporting a high need for affiliation whereas those cultures that have been classified as masculine tend to have a higher number of people with a high need for achievement.

According to David McClelland's Acquired Needs Theory (also called Learned Needs Theory), individuals are motivated by needs that are socially acquired or learned from their culture (McClelland, 1961). McClelland's research identified three types of needs that motivate individuals. The three needs that motivate workers include the need for achievement, the need for affiliation, and the need for power. Adults are assumed to possess all three categories of needs to one degree or another; however, one of the motives is usually dominant.

Achievement orientation is "a concern for performing well or for striving for a standard of excellence" (The Center for Leadership Excellence, 1995: 20). Individuals with a high need for achievement behave in ways to improve job performance and strive to do tasks faster and more efficiently (Spencer, McClelland, & Spencer, 1994). They focus on personal achievements rather than on the rewards of success. They set challenging goals for themselves and are continuously focused on finding ways to do things better (The Center for Leadership Excellence, 1995). This focus on continuous improvement could lead to higher quality goods and services and increased organizational productivity (Spencer et al., 1994). Companies with a large number of employees with a high need for achievement grow faster and are more profitable.

The need for affiliation is characterized by a desire to belong, an enjoyment of teamwork, a desire for friendly and close interpersonal relationships, and a need to reduce uncertainty. These individuals are interested in social relationships and prefer cooperative rather than competitive work environments. As managers, highly affiliative individuals may be reluctant to monitor the behavior of subordinates, give negative feedback to others, or discipline their subordinates.

The need for power is characterized by a drive to control and influence others, a need to win arguments, a need to persuade and prevail. Individuals motivated by power are concerned about their impact on other people - convincing someone of their point of view or empowering others around them, and finding ways to connect with and influence powerful people.

Given the relationship between gender-role orientation and McClelland's needs, it is pertinent to include them in our study. As stated earlier, the need for affiliation is considered to be feminine in nature. In addition, it has been found that the need for achievement is considered to be masculine in nature and the need for power, by definition, should also be more masculine.

METHODOLOGY

In order to measure the dimension of masculinity-femininity and the related values of the motivating needs for achievement and affiliation, surveys were administered to both female and male students at a large university in Seoul, South Korea. In addition, general societal values were also measured. The overall survey included a detailed participant data sheet to collect information on demographic variables (i.e., sex, age, ethnicity, international influences throughout life, and work experience) and questionnaires to assess general values (Rokeach Value Scale or RVS), gender-role orientations (Bem Scale Short Form or BSRI), and motivating needs (Personal Values Questionnaire or PVQ).

Gender-Role Orientation Questionnaire – Bem Sex Role Inventory

The Bem Sex-Role Inventory (Short Form) published by Consulting Psychologists Press, Inc. was administered to categorize individuals as masculine, feminine, androgynous, or undifferentiated in their gender-role orientations.

Motivating Needs of Achievement and Affiliation – Personal Values Survey

The need for achievement, the need for affiliation, and the need for power were measured through the Personal Values Questionnaire, which was developed by McClelland for McBer and Company (1993). Ten questions relate to the need for affiliation, ten questions relate to the need for achievement, and ten questions relate to the need for power.

Societal Values – Rokeach Value Survey

In addition, the Rokeach Value Survey (RVS) was administered to get an overview of the cultural values present in South Korea. Rokeach (1973) designed this instrument to measure personal and social values. The RVS distinguishes two types of values, instrumental and terminal. Instrumental values refer to preferred modes of behavior while terminal values refer to desired end states of existence.

At this point, the data has been collected but not fully analyzed. Preliminary results seem promising and we believe that the completed manuscript will provide valuable information. There is a need for more research on South Korean cultural and personal values and the implications related to them. There is much evidence that different personal, cultural, and work values cause communicative challenges and misunderstandings. Different value systems often manifest themselves in different work habits and expectations from an employer/employee relationship. Decision making styles, motivation techniques, leadership behaviors, etc. will all be influenced by the value systems of those involved. This research seeks to further the understanding of the South Korean value system and will prove useful for practicing US managers in the development of organizational strategies and training programs.

References Available Upon Request

China: World Trade Organization and Trade Disputes

ABSTRACT

Since China's membership in the World Trade Organization (WTO), the Chinese government has had to adjust to the WTO rules and regulations in regard to the fair trade practices. The Chinese government has made progress in changing its traditional trade practices; however, the government faces a tremendous challenge in implementing the necessary market reforms. As a result, China has encountered major complaints and trade disputes. This research examines the trade disputes that have been brought to the WTO by China or against China. The disputes are studied focusing on some important industries, such as agriculture, electronics, and textile.

FINANCIAL STATEMENT ANALYSIS USING DATA ENVELOPMENT ANALYSIS

D.K. Malhotra, Philadelphia University, (215) 951-2813, MalhotraD@philau.edu
Rashmi Malhotra, St. Joseph's University, (610) 660-3497, rmalhotr@sju.edu

ABSTRACT

This article explains how data envelopment analysis can be used as a decision support system to perform financial statement analysis and demonstrates the benefits DEA offers over traditional financial statement analysis techniques. A major advantage of the DEA approach is that it clearly identifies the factors contributing to the performance of company over its competitors.

INTRODUCTION

Financial statements are a summary of the operating, financing, and investment activities of a firm over a period of time. Financial statements are supposed to contain enough information to help investors and creditors make an informed decision about investing or lending money to the company. Financial statement analysis is the ultimate key that will help investors and creditors gain enough insight into the company to make an informed decision about the company. Financial statement analysis is also used by management to make decisions about the firm in a more informed manner. Financial statement analysis helps identify a firm's strengths and weaknesses so that management can take advantage of a firm's strengths and make plans to counter weaknesses. The strengths must be understood if they are to be used to proper advantage and weaknesses must be recognized if corrective action needs to be taken. From management's viewpoint, financial statement analysis is useful both as a way to anticipate future conditions, starting point for planning actions that will influence the future course of events or to show whether a firm's position has been improving or deteriorating over time.

Ratio analysis is an integral part of financial statement analysis. Ratio analysis begins with the calculation of a set of financial ratios from the information contained in the financial statements (income statement, balance sheet, statement of cash flows). Ratios analysis is designed to show the relative strengths and weaknesses of a company as compared to

- Other firms in the industry
- Leadings firms in the industry, and
- the previous year of the same firm.

Ratio analysis helps to show whether the firm's position has been improving or deteriorating relative to its own past or other firms in the industry or industry averages. By benchmarking the firm's financials against its own peers or industry averages, management can identify the relative strengths and weaknesses of the firm and plan better for the future. Benchmarking will also help investors and creditors better understand the relative position of a firm in the industry and make investment/lending decisions in a more informed manner.

In this paper, we illustrate the use of data envelopment analysis, an operations research technique, to analyze financial statements of firms by benchmarking financial ratios of a firm against its peers as well as against the industry averages. We use the pharmaceutical industry as an example in this study to

conduct our analysis. We show the advantages of data envelopment analysis in benchmarking a firm. Data envelopment analysis also helps identify the areas in which a firm has strengths relative to competition as well as the areas in which the firm is weak relative to competition. Furthermore, data envelopment analysis also lays the areas in which a firm needs improvement relative to its peers and by how much improvement is needed in each of those areas.

The rest of the paper is organized along the following lines. In section II, we provide a review of previous studies that illustrate the use of data envelopment analysis in the area of business. Section III discusses data envelopment analysis and the model used in this study. In section IV, we provide a discussion of empirical results, and section V summarizes and concludes the study.

PREVIOUS STUDIES

A number studies have been published on different aspects of financial statement analysis. We include only those studies that use data envelopment analysis in either financial statement analysis or analysis of financial performance of firms. Zhu (2000) uses data envelopment analysis to develop a multi-factor financial performance model that recognizes tradeoffs among various financial measures. Kao and Liu (2004) compute efficiency scores based on the data contained in the financial statements of Taiwanese banks. They use this data to make advanced predictions of the performances of 24 commercial banks in Taiwan. Pille and Paradi (2002) analyze the financial performance of Ontario credit unions. They develop models to detect weaknesses in Credit Unions in Ontario, Canada. Oscan and McCue (1996) use data envelopment analysis for measuring and assessing the financial performance for hospitals. They compute a financial performance index (FPI) as a measure of aggregate financial performance. They show that financial performance index across many financial ratios eases the comparison of an individual hospital with its peers. Feroz, Kim and Raab (2003) is the only study that directly talks about financial statement analysis using data envelopment analysis methodology. They show that data envelopment analysis can augment the traditional ratio analysis to a consistent and reliable measure of managerial or operational efficiency of a firm.

METHODOLOGY

WHAT IS DATA ENVELOPMENT ANALYSIS?

Data envelopment analysis is a technique used to assess the productive efficiency of homogenous operating units such as schools, hospitals, banks, or utility companies. It is a powerful technique for measuring performance because of its objectivity and ability to handle multiple inputs and outputs that can be measured in different units. The DEA approach does not require specification of any functional relationship between inputs and outputs or a priori specification of weights of inputs and outputs. DEA provides gross efficiency scores based on the effect of controllable and uncontrollable factors. DEA uses a number of financial ratios to determine how good a company's performance has been. A firm's performance is analyzed on the basis of a set of financial ratios that include liquidity ratios (current ratio, quick/acid test ratio), asset management ratios (inventory turnover ratio, asset turnover ratio), debt management ratios (leverage ratio, total debt to total assets, and times interest covered ratio), and profitability ratios (Return on Equity, Return on Assets, and net profit margin). The ratios that need to be maximized serve as outputs and ratios that need to be minimized serve as inputs. Using this information, approach does not require specification of any functional relationship between inputs and outputs or a

priori specification of weights of inputs and outputs. DEA provides gross efficiency scores based on the effect of controllable and uncontrollable factors. DEA uses a number of variables to determine how good a firm is. With these financial ratios as inputs, the DEA-based decision support system calculates 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 either an existing firm or a hypothetical firm with an identical set of inputs or outputs that is constructed as a combination of good 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. This helps investors and creditors justify their decisions to invest or not to invest their funds in a particular company. This will also help management identify areas of weakness for a firm so that management plans can focus on plugging the weaknesses or taking steps to counter the weaknesses.

USING DEA FOR FINANCIAL STATEMENT ANALYSIS

Traditional financial statement analysis techniques use ratio analysis to compare a firm’s performance against its peers in the industry as well as against the company’s historical performance. On the basis of this comparison, analyst will recommend whether the company is doing well or underperforming relative to its peers or relative to its own past performance. 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. The higher the current ratio or quick ratio or inventory turnover ratio or asset turnover ratio, the higher the chances of a firm turning out to be efficient. Similarly, the higher the leverage ratio or total debt to total capital ratios, the higher the chances that the firm will turn out to be bad. The acceptance boundary has firms that are considered to be 100% efficient. Ratio analysis is a commonly used analytical tool for verifying the performance of a firm. While ratios are easy to compute, which in part explains their wide appeal, their interpretation is problematic, especially when two or more ratios provide conflicting signals. Indeed, ratio analysis is often criticized on the grounds of subjectivity, that is the analyst must pick and choose ratios in order to assess the overall performance of a firm. In this paper we demonstrate that Data Envelopment Analysis (DEA) can augment the traditional ratio analysis. DEA can provide a consistent and reliable measure of managerial or operational efficiency of a firm.

EMPIRICAL ANALYSIS

In order to evaluate the effectiveness of DEA to benchmark companies, we collect data on sixteen pharmaceutical companies. We use days of sales outstanding, days cost of goods sold in inventory, total debt/equity as input variables. On the other hand, we use cash flow per share, return on equity, return on assets, return on invested capital, inventory turnover, asset turnover, current ratio, quick ratio, and interest rate coverage as output variables. The DEA model maximizes the output variables and minimizes the input variable to compare the relative performance of different companies. To benchmark companies, we consider each of the companies as a homogenous unit, and we can apply the DEA methodology to assess a comparative performance of these companies. Using the DEA methodology, we can calculate an efficiency score for the 16 companies on a scale of 1 to 100. Table 1

illustrates the efficiency scores of the 16 companies. Further, we also study the peers (model companies) for inefficient companies.

Table 1 illustrates that eleven out of sixteen companies are fully efficient. Furthermore, table 2 presents the efficiency rankings in ascending order. We present the score in percentage value varying between 0% and 100%. We find that the output efficiency of Bayer AG, Abbott, Mylan, AMGN, Schering-Plough, Merck, Bausch and Lomb, Pfizer, Glaxo Smithkline, Johnson and Johnson, and Wyeth is 100%. On the other hand, the output efficiency of the remaining companies are: Baxter – 97% , Novartis - 92%, Novartis – 92%, Elli Lilly – 88%, Teva – 82%, and Bristol Myers Squibb – 77%. This means that the observed levels of cash flow per share, return on equity, return on assets, return on invested capital, inventory turnover, asset turnover, current ratio, quick ratio, and interest rate coverage are .97 times the maximum output level that Baxter can secure with its current days of sales outstanding, days cost of goods sold in inventory, total debt/equity. The same rationale applies to Elli Lilly, Novartis, Teva, and Bristol Meyers Squibb. The remaining eleven companies turn out to be the best practices.

The best practices companies: Bayer AG, Abbott, Mylan, AMGN, Schering-Plough, Merck, Bausch and Lomb, Pfizer, Glaxo Smithkline, Johnson and Johnson, and Wyeth are 100% efficient. The remaining companies are inefficient. Therefore, 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.

Table 2 illustrates the peer group for the inefficient companies. As shown in the Table 2, Schering-Plough, Johnson and Johnson, and Bausch and Lomb serve as peer for Teva. In addition, Teva is more comparable to Bausch and Lomb (weight 47%) and less comparable to its more distant peer Johnson and Johnson (weight 33%), and further distant peer Schering-Plough (weight 20%). Thus, Teva should scale down its levels of current days of sales outstanding, days cost of goods sold in inventory, total debt/equity other factors to make them comparable with Bausch and Lomb. Similarly, Merck, Wyeth, Schering-Plough, Pfizer, Bayer AG, and AMGN serve as peers for other companies.

After calculating the efficiency of a company using DEA, and identifying the efficient peers, the next step in DEA analysis is feasible expansion of the output or contraction of the input levels of the company within the possible set of input-output levels. The DEA efficiency measure tells us whether or not a given company can improve its performance relative to the set of companies to which it is being compared. Therefore, after maximizing the output efficiency, the next stage involves calculating the optimal set of slack values with an assurance that output efficiency will not increase at the expense of slack values of the input and output factors. Once efficiency has been maximized, the model does seek the maximum sum of the input and output slacks. If any of these values is positive at the optimal solution to the DEA model that implies that the corresponding output of the company (DMU) can improve further after its output levels have been raised by the efficiency factor, without the need for additional input. If the efficiency is 100% and the slack variables are zero, then the output levels of a company cannot be expanded jointly or individually without raising its input level. Further, its input level cannot be lowered given its output levels. Thus, the companies are pareto-efficient with technical output efficiency of 1. If the company is 100% efficient but one slack value is positive at the optimal solution then the DEA model has identified a point on the efficiency frontier that offers the same level on one of the outputs as company A in question, but it offers in excess of the company A on the output corresponding to the positive slack. Thus, company A is not Pareto-efficient, but with radial efficiency of 1 as its output cannot be expanded jointly. Finally, if the company A is not efficient ($<100\%$) or the efficiency factor is greater than 1, then the company in question is not Pareto-efficient and efficiency

factor is the maximum factor by which both its observed output levels can be expanded without the need to raise its output. If at the optimal solution, we have not only output efficiency > 1 , but also some positive slack, then the output of company A corresponding to the positive slack can be raised by more than the factor output efficiency, without the need for additional input. The potential additional output at company A is not reflected in its efficiency measure because the additional output does not apply across all output dimensions. Table 3 illustrates the slack values identified in the next stage of the DEA analysis. The slack variables for 100% efficient companies Bayer AG, Abbott, Mylan, AMGN, Schering-Plough, Merck, Bausch and Lomb, Pfizer, Glaxo Smithkline, Johnson and Johnson, and Wyeth.

These companies 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. However, the slack variables are non zero for Baxter, Elli Lilly, Novartis, Teva, and Bristol Meyers Squibb. As illustrated in the table, for Teva, besides increasing the output levels of interest rate coverage by 28.73 units, should reduce the days of sales outstanding by 26.05 units. Teva should follow Schering Plough, Johnson and Johnson, Bausch and Lomb as role models. Similarly, we can interpret the slack variables for Elli Lilly, Novartis, Teva, and Bristol Meyers Squibb.

The next step in our analysis is to perform sensitivity analysis of the DEA model. DEA is an extreme point technique because the efficiency frontier is formed by actual performance of best-performing DMUs (Ramanathan 2003). Furthermore, as DEA is a non-parametric technique, statistical hypothesis tests are difficult. It is possible for a DMU to obtain a value of utility by simply improving its performance in terms of only one particular output ignoring others. One way of checking the sensitivity of DEA efficiency of a DMU is by omitting one or more inputs or outputs. Thus, we used 18 different models to calculate efficiency of the pharmaceutical companies. Table 4 summarizes the results of our analysis. Table 4 displays the average efficiency, the standard deviation of the efficiencies, and median efficiency level for each country. Table 5 lists all the countries and their rankings based on average efficiency. As expected, Johnson and Johnson, Merck, Abbott, Bausch and Lomb, and AMGN are the most efficient.

SUMMARY AND CONCLUSIONS

This study illustrates the use of data envelopment analysis to analyze financial statements of firms. We compare the relative performance of the 16 pharmaceutical companies using financial ratios. We use days of sales outstanding, days cost of goods sold in inventory, total debt/equity as input variables. On the other hand, we use cash flow per share, return on equity, return on assets, return on invested capital, inventory turnover, asset turnover, current ratio, quick ratio, and interest rate coverage as output variables. The DEA methodology benchmarks best-performing companies against worst-performing companies. Using the DEA methodology, we compute an efficiency score for the 16 companies on a scale of 1 to 100. We find that ten companies are efficient with a score of 100 and six companies show a score below 100 and, therefore, are inefficient. By computing slack variables, we also show the areas in which inefficient firms need to improve. By providing an overall efficiency score, data envelopment analysis eliminates the need to interpret conflicting ratios. Through slack variables, we are also able to identify the areas in which the firm needs improvement.

TABLES & REFERENCES

Tables, references, and full paper available upon request from the authors.

ANALYZING THE RELATIONSHIPS BETWEEN QUALITY, TIME, AND COST IN PROJECT MANAGEMENT DECISION MAKING

Matthew J. Liberatore, Department of Management and Operations, Villanova University, Villanova, PA 19085, 610-519-4390, matthew.liberatore@villanova.edu

Bruce Pollack-Johnson, Department of Mathematical Sciences, Villanova University, Villanova, PA 19085, 610-519-6926, bruce.polack-johnson@villanova.edu

ABSTRACT

This paper presents a mathematical programming model that allows quality to be explicitly considered in project planning and scheduling, while addressing the tradeoffs between quality, time, and cost. Using a construction example we show how this model can be used to generate quality level curves to illustrate the trade-offs among time, cost, and quality. These level curves can then be used by project managers to make project scheduling decisions that explicitly model and consider quality as well as time and cost, so that better and more appropriate decisions can be made for a particular situation.

Keywords: project management, quality management, project planning, project scheduling, mathematical programming

INTRODUCTION

Project management requires achieving cost, schedule, and performance targets, while providing an outcome that satisfies the client. A measure of the value of the project to the client is the level of quality associated with the completed project. It follows that important managerial decisions relate to the level of quality achieved for each of the project's tasks, since *in toto* the quality of the tasks defines the quality of the project.

The emphasis in project planning and scheduling has been on the relationship between time and cost, with little attention directed to quality. In most situations there are alternate approaches for completing each task, each having its own time, cost, and quality. The managerial question of interest relates to these choices, which have profound impacts on the project outcome.

The purpose of this paper is to present a modeling framework that allows quality to be explicitly considered in project planning and scheduling, while addressing the tradeoffs between quality, time, and cost. This paper also offers some managerial insights that are derived from the modeling framework through improved understanding of these choices and tradeoffs.

PROJECT QUALITY

Quality can be defined as “a dynamic state associated with products, services, people, processes, and environments that meet or exceed specifications” [2]. The ISO 9000 definition of quality [3] has been adopted by the Project Management Body of Knowledge [7]. According to the PMBOK, quality must address both the management of the project and the product of the

project. Our interest in quality relates to the quality planning process, and addresses quality at the task and overall project levels.

Paquin et al. [5] contend that a method for assessing quality must enable project managers to elucidate and structure the client's needs and expectations. We are interested in measuring planned quality of work for different designs of specific activities using such a method. We assume that there are choices for completing tasks that vary in quality, time, and cost.

The suggested approach requires the identification of the quality attributes that are relevant for the project. An example using the Analytic Hierarchy Process (AHP) [8] [9] to evaluate the quality of a task option along with further discussion of this approach can be found in [6].

MODELING FRAMEWORK

We begin by formulating a model of the quality of each individual task as a function of the time and cost allocated to it. We assume that there could be different entities who could do the task and that each entity could do the job with different allocations of time and budget. Each entity would have its own quality function in terms of time and cost. If those quality functions are graphed on the same time/cost/quality axes, then the overall quality function for the task that we are interested in is the *maximum*, or the *upper envelope*, of the individual entity quality graphs.

We assume that this overall quality function for a task has two basic properties:

- Holding time constant, quality is *nondecreasing* in cost. Thus if time is fixed, we assume that spending more money on the task will *increase* (or at least not decrease) the quality.
- Holding cost constant, quality is *nondecreasing* in time. Thus if cost is fixed, we assume that spending more time on the task will *increase* (or at least not decrease) the quality.

If we normalize quality to be on a 0-100 scale, and limit time and cost to reasonable values for the task at hand, based on the two nondecreasing assumptions above, we would expect the graph to show the quality being lowest at the corner of the domain with the smallest values of time and cost and highest in the opposite corner (the highest values of time and cost). For a fixed quality, we would expect a traditional time/cost tradeoff curve, which is normally a decreasing convex curve (to maintain the same level of quality, to reduce the time, one has to pay more money, such as in standard project activity *crashing* [1]). This suggests a basic hill shape rising out of a plain, although we would only be interested in a one-quarter wedge of the hill.

A familiar mathematical functional form that has this shape is the bivariate normal distribution* in probability. We propose using this functional form for the overall quality function for each task. Our quality function is normalized so that the maximum time (μ_t) and cost (μ_c) values considered reasonably correspond to a quality of 100. The "standard deviation" parameters (σ_t and σ_c) give a measure of how slowly the quality drops from the top of the hill compared to the maximum values for time and cost, respectively. Thus, our resulting quality function is given by

* This version of the bivariate normal distribution assumes independence of the two random variables. We have chosen this version for a simpler model, but the dependent version could also be used, with one more parameter, corresponding to the correlation between the variables.

$$Q(t, c) = 100e^{-\left[\left(\frac{t-\mu_t}{\sigma_t}\right)^2 + \left(\frac{c-\mu_c}{\sigma_c}\right)^2\right]}$$

We have normalized the initial constant to 100 and eliminated the $\frac{1}{2}$ in the exponent (which means that each σ would be multiplied by $\sqrt{2}$ to be interpreted as the usual σ in the bivariate normal). If we hold either variable constant, the marginal graph for the other will be a bell curve (actually, a *subset* of a graph that is a constant multiple of a normal distribution curve). The upper envelope graph may not be smooth, but we are assuming that we can create a smooth function that is a reasonable estimate of the upper envelope.

In situations where n bids specifying levels of quality, time, and cost (q_j, t_j, c_j) have been received for a given activity, the four parameters of the bivariate normal function can be determined using nonlinear least squares estimation.

MODEL FORMULATION

We start with standard assumptions for modeling projects: that the project network has no cycles, that the start activity (activity 0, a dummy activity) is the only activity that is not an immediate successor¹ of any activity, and that the finish activity (activity $N+1$, also a dummy activity) is the only activity that has no successors. Define the following parameters and variables:

- t_i = the duration of activity i , for $i = 1, \dots, N$
- c_i = the cost of activity i , for $i = 1, \dots, N$
- q_i = the quality of activity i , for $i = 1, \dots, N$
- S_i = the set of activities that are immediate successors of activity i , for $i = 0, \dots, N$
- T_{UB} = upper bound on the total project time
- Q_{LB} = lower bound on project quality
- st_i = the scheduled start time for activity i , for $i = 0, \dots, N+1$
- t_{\min_i} = lower bound on the duration of activity i , for $i = 1, \dots, N$
- c_{\min_i} = lower bound on the cost of activity i , for $i = 1, \dots, N$

Relevant quality measures could involve maximizing average quality, or maximizing minimum quality, of the tasks. We select the latter, Q_{\min} , as our quality metric, since from a systems perspective if the project is viewed as an integrated set of activities, the quality of a project is only as high as its weakest link. Q_{\min} is defined as:

$$Q_{\min} = \min_{1 \leq i \leq N} q_i \tag{1}$$

In our formulation, we minimize total project cost while setting a lower bound on Q_{\min} and an upper bound on total project time. The nonlinear program is given as equations (2) – (12):

¹ It is common to use predecessors rather than successors for formulations of this type, but for this example, the formulation turns out to be much more concise and elegant using successors.

$$\text{Minimize } \sum_{i=1}^N c_i \quad (2)$$

Subject to:

$$Q_{\min} \leq q_i, i = 1, 2, \dots, N \quad (3)$$

$$Q(t_i, c_i) = 100 * \exp \left\{ -\left[\frac{(t_i - \mu_{t_i})}{\sigma_{t_i}} \right]^2 - \left[\frac{(c_i - \mu_{c_i})}{\sigma_{c_i}} \right]^2 \right\}, i = 1, 2, \dots, N \quad (4)$$

$$Q_{\min} \geq Q_{LB} \quad (5)$$

$$st_0 = 0 \quad (6)$$

$$st_k \geq st_i + t_i \quad \forall i = 0, \dots, N, \forall k \in S_i \quad (7)$$

$$st_{N+1} \leq T_{UB} \quad (8)$$

$$st_i \geq 0 \quad \forall i = 1, \dots, N + 1 \quad (9)$$

$$t_{\min_i} \leq t_i \leq \mu_{t_i}, i = 1, 2, \dots, N \quad (10)$$

$$c_{\min_i} \leq c_i \leq \mu_{c_i}, i = 1, 2, \dots, N \quad (11)$$

$$q_i, t_i, c_i \geq 0, i = 1, 2, \dots, N \quad (12)$$

This problem can be solved using Lingo's global solver [4] and extends the standard cost – time tradeoff problem [1].

CONSTRUCTION EXAMPLE

A general contractor planning to start construction of a new house has organized the project into activities as given in Table 1. The corresponding project network diagram is shown in Figure 1. She has received bids for both duration and cost from different subcontractors. These bids were used to estimate the bivariate normal quality functions for each activity (Table 1).

USING QUALITY LEVEL CURVE GRAPHS

One way to evaluate the interactive relation among project time, total cost, and quality is to create level iso-quality graphs. Specifying a value of Q_{LB} , for different total project times (upper limits), using our model we can then find the minimum cost possible that finishes the project within a given time and maintains a minimum quality of at least Q_{LB} . A set of level iso-quality curves for the construction example is shown in Figure 2.

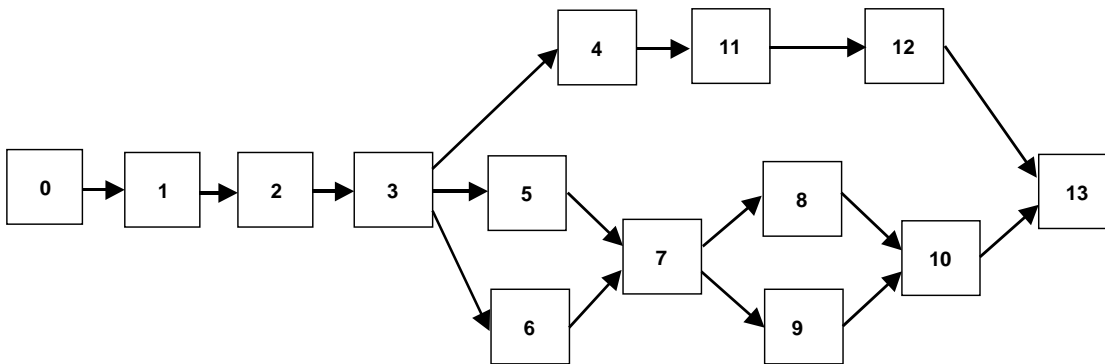
The graph for a higher quality level lies above and to the right of that for a lower quality level, although they can overlap in places for quality levels that are very close together. There are several places where a level curve is horizontal. This could mean that to achieve a certain quality level, a choice may need to be made at a *longer* project time value that forces a solution which actually finishes the project in strictly *less* than the upper limit for the total time, and therefore the same solution is optimal at a *shorter* project time limit. Figure 2 provides a concise summary of the relationship among time, cost, and quality, and can be used to make well-informed decisions about how to execute the project.

Table 1: Task, Immediate Successor, and Quality Function Information for Construction Project

TASK	DESCRIPTION	IMMEDIATE SUCCESSORS	QUALITY PARAMETERS ($\mu_t, \sigma_t, \mu_c, \sigma_c$)
0	START	1	dummy activity
1	Excavate and Pour Footers	2	Not estimated – one bid*
2	Pour Concrete Foundation	3	Not estimated – one bid*
3	Erect Rough Wall & Roof	4,5,6	(4, 1.79, 48.6, 42)
4	Install Siding	11	(13, 19, 79.2, 99.4)
5	Install Plumbing	7	(3, 1.62, 26.6, 20.4)
6	Install Electrical	7	(10.9, 12.8, 29.7, 77.9)
7	Install Wallboard	8,9	(5, 2.73, 16.8, 8.05)
8	Lay Flooring	10	(8.09, 7.18, 64, 67.5)
9	Do Interior Painting	10	(4.57, 4.3, 16.8, 12.7)
10	Install Interior Fixtures	13	Not estimated – one bid*
11	Install Gutters & Downspouts	12	(2, 12, 17.7, 18.9)
12	Do Grading & Landscaping	13	(3.36, 2.4, 21.5, 12.3)
13	FINISH	--	dummy activity

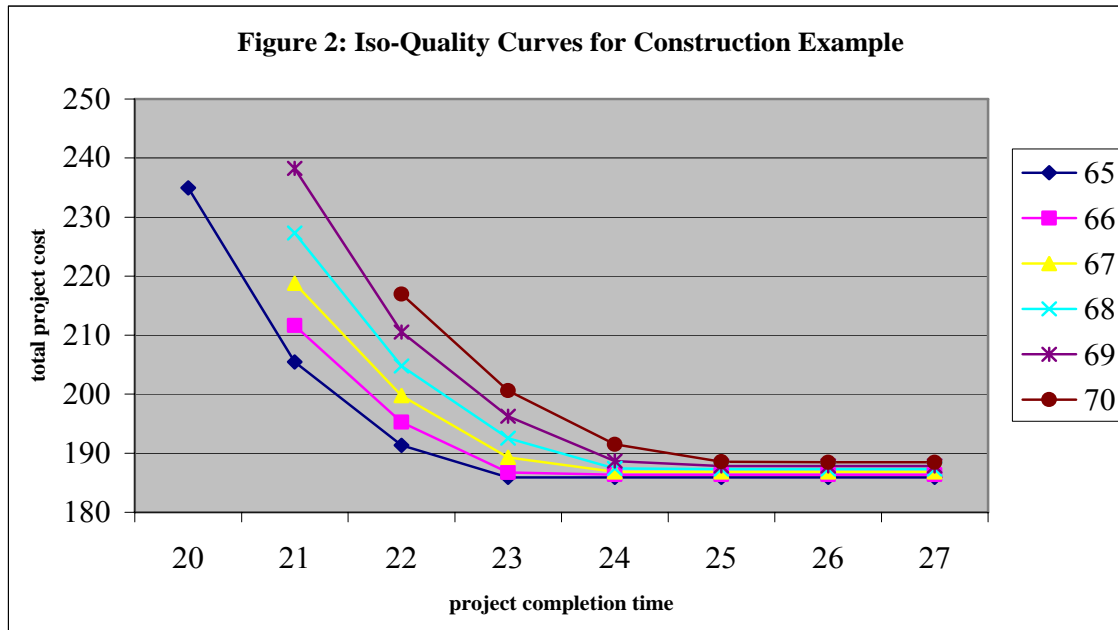
*for those activities having one bid, the quality, time, and cost (q, t, c) estimates were used directly in the analysis: activity 1: (70, 3, 26.6); activity 2: (70, 1, 7.2); activity 10: (70, 3, 7.2)

Figure 1: Project Network Diagram for Construction Example



CONCLUSIONS

In standard project planning and scheduling, quality is acknowledged to be important at different levels, but previously has not been explicitly modeled. In many situations there are alternate options for accomplishing project activities, and these involve differing levels of time, cost, and quality. In such situations it makes sense to model the relationship between cost, time, and quality, and determine their levels for each activity that best achieves the project's objectives. We have presented a nonlinear programming model for the quality/time/cost problem, and have shown how quality level curves can be a very useful management tool in making final project scheduling decisions that explicitly model and incorporate quality.



REFERENCES

- [1] Brucker, P., Drexl, A., Mohring, R., Neumann, K., Pesch, E. "Resource-constrained project scheduling: Notation, classification, models, and methods," *European Journal of Operational Research*, 1999, 112(1), 3- 41.
- [2] Goetsch, D. L., Davis, S. B. *Quality management* (5th ed.). Upper Saddle River, NJ: Pearson Prentice-Hall, 2006.
- [3] International Organization for Standards, *ISO 9000:2000*, 2000.
- [4] Lindo Systems. *Lingo Version 9.0*. Chicago, 2004.
- [5] Paquin, J. P., Couillard, J., Ferrand, D. J. "Assessing and controlling the quality of a project end product: The earned quality method," *IEEE Transactions on Engineering Management*, 2000, 47(10), 88-97.
- [6] Pollack-Johnson, B., Liberatore, M. "Incorporating Quality Considerations into Project Time/Cost Trade-off Analysis and Decision Making," *IEEE Transactions on Engineering Management*, 2006, 53(4), 534 – 542.
- [7] Project Management Institute. *A Guide to the project management body of knowledge* 3rd ed. Newtown Square, PA, 2004.
- [8] Saaty, T. L. "A Scaling method for priorities in hierarchical structures," *Journal of Mathematical Psychology*, 1977, 15, 234-81.
- [9] Saaty, T. L. *The analytic hierarchy process*. Pittsburgh: RWS Publications, 1996.

BILEVEL OPTIMIZATION OF BIOFUEL PRODUCTION WITH FIXED COST

Xiangrong Liu, Drexel University, PA 19104, USA, xl35@drexel.edu
Hande Benson, Drexel University, PA 19104, USA, hvb22@drexel.edu

ABSTRACT

Bilevel optimization problems consist of an upper level optimization problem whose sets of feasible solutions depend on the set of optimal solutions of a lower level problem. The application considered is biofuel production. In the upper level problem, government finds the optimal tax exemption level to induce industry and farms to produce biofuel. In the lower level problem, farms allocate their lands to different usages in response to the government policy. We focus on bilevel optimization problems where the lower level problems have binary variables. A numerical example is presented and solved using grid search method.

Keywords: bilevel optimization, discrete variable, biofuel production, fixed cost

INTRODUCTION

Biofuel is a type of fuel derived from biomass and used for power and automotive transport. Many crops such as wheat, corn, rapeseed and sunflower can be the resources to produce biofuel. Recently, with more concerns over rising oil prices, gas emission effects and rural development interests significantly drive the biofuel use and development. The legislation, HB3543 signed by Oregon Governor Ted Kulongoski in July 2007, requires “all gasoline sold in the state to be blended with 10% bioethanol and all diesel fuel sold in the state to be blended with 2% biodiesel” [14]. In order to break its dependence on imported oil, the U.S. government now provides more funding, including \$179 million for the President’s Biofuels Initiative[9] and \$ 375 million by the Department of Energy[13], to reduce the cost and improve the efficiency of biofuel production.

Government, industry and farms form a supply chain through which material, product and currency flow. Generally, there are three levels of decision making in this biofuel production problem: 1)As a leader, the government plays an active role in encouraging biofuel production. Under economical, environmental and social considerations, government would like to provide direct incentives such as tax exemptions to industry. Naturally, the cost to the government is expected to be minimized while ensuring the required amounts of production; 2)Due to the high production cost of biofuel, industry expects more tax exemption from government to ensure its own profit. Otherwise, industry has no incentive to replace oil with biofuel products. The industry will determine the price paid to the farmers at the gate; 3)Because of surging demand for biofuels, farms benefit by switching from producing food crops to nonfood crops, which are the raw materials for biofuel production. Farms choose among three options: planting food crops, growing nonfood crops or leaving a part of the land fallow. The strategic plan needs to be set for the allowable land to be better utilized to realize the maximum profits.

Models involving such decisions for biofuel production are typically formulated as bilevel programming problems. At each level, a certain objective function needs to be optimized without forcing lower levels to make suboptimal decision. 1)In the upper level problem, the government needs to minimize its total expenditure by setting the optimal tax exemption level provided to the industry. Meanwhile, the industry will determine the price paid to the farmers at the gate and pass the incentives onto the farmers. 2)In the lower level problem, farms decide the optimal assignments of allowable lands to maximize their profits.

However, in order to ensure sufficient food and enough fuel, a license fee might be charged if the farms sell crops as a biofuel resource. In addition, transportation and equipment costs for nonfood crops may be incurred because of the new destination and purpose of the final products. We category all these costs into fixed costs. As a result, the increased importance of the biofuel development program with fixed cost stimulated our interest to improve modeling and solution methods used in the past. The introduction of binary variables is essential in the agronomic practice and policy making, despite adding significant complexity to the computation of optimal solution.

The basic framework of the paper is organized as follows: in the next section, we review the literature related to biofuel production and discrete bilevel optimization; Next, we introduce our new model based on that of Bard et al.[3] and present a method to solve the bilevel problem with discrete variables in the lower level. Finally, in the last section, we provide a numerical example to show the implications of incorporating fixed costs in the model.

LITERATURE REVIEW

For a thorough treatment of the development and foundations of bilevel programming, we refer interested readers to the books by Bard [1] and Dempe [7]. More recently, Colson [5], Dempe [8] and Vicente [17] have provided thorough surveys of current research in the field. Bilevel programming can be applied in many different contexts such as transportation and resource allocation planning [12]. However, bilevel models with discrete variables have remained relatively unstudied in literature. Bard et al.[2], Dempe [6] and Vicente [18] provide algorithms for solving discrete binary bilevel programming problems.

The majority of the biofuel literature focuses on the political decisions and economic cost and benefit analysis of biofuel production according to the practice in different countries. Rozakis et al. [16] use a partial-equilibrium model to discuss the biofuel cost of chains operating under different policy scenarios in France. The possibilities of biofuel cost reduction are analyzed. Monte Carlo simulation is implemented to deal with uncertainty and minimal tax exemption levels for the viability of the activity are determined. The study by Kruse et al. [10] uses a stochastic model to find out the effects of extension of the tax credits and import tariff on future growth in biofuels. The trade-off between maximizing farm income and minimizing government costs are balanced by setting up the appropriate tax credits and import tariff structure. Rajagopal and Zilberman [15] review the literature on the biofuel problem from the perspective of environment, economics and policy-setting.

Regarding the application of bilevel programming to the biofuel problem, the study by Candler et al. [4] points out the potentials and the difficulties. Bard et al [3] formulate a leader-follower game that helps decision makers arrive at a rational policy for encouraging biofuel production. In this paper, we base our work on [3] and incorporate fixed costs in the farm's problem to further extend the model. Doing so requires that we include binary variables in the lower level, which greatly increases the computational complexity. However, the resulting model is more consistent with the practice.

MODELS

Notation And Assumption

We follow the notation of Bard et al[3]. Meanwhile, all the farms are assumed to have the same set of plants that could be produced in the arable lands.

- Parameters

- α_{db} = conversion rate from nonfood crops d to biofuel b (FF/hl)
- r_{df} = yield rate of nonfood crop d grow on farm f (t/ha)
- γ = set-aside payment for fallow land (FF/ha)
- η_d = indicator parameter: = 1 if no subsidy is paid for nonfood crop d
- ρ = maximum percentage of set-aside land expected by government for nonfood crops
- θ = percentage of arable land required to be set aside for nonfood crops or fallow
- w_f = total arable land available on farmer f (ha)
- u_b = upper bounds on production of biofuel b from certain nonfood crops (ha)
- π_b = profit expected by industry for one unit of biofuel b (FF/hl)
- v_b = market price for biofuel b (FF/hl)
- β_{db} = cost of converting one tone of nonfood crops d to biofuel b (FF/ha)
- o_{db} = co-products price associated with one unit of biofuel b producing from nonfood crop d
- m_{cf} = gross margin for food crop c grown on farm f (FF/ha)
- s_d = subsidy paid to farms for nonfood crop d (FF/ha)
- c_{df} = net production cost for nonfood crop d on farm f (FF/ha)
- κ_{df} = upper bounds on percentage of land permitted for crops in agronomic constraint
- δ = maximum percentage of land expected by government for fallow
- t_{df} = fixed cost for farm f for planting nonfood crop d

- Sets

- B = set of biofuels
- F = set of farms
- C = set of food crops
- D = set of nonfood crops
- $D(b)$ = subset of nonfood crops used to make biofuel b

- Decision Variables

- τ_b = government tax credit given to industry for biofuel b (FF/ha)
- p_d = price paid by industry to farms for nonfood crop d (FF/t)
- x_{cf} = area allocated to food crop c by farm f (ha)
- xn_{df} = area allocated to food crop d by farm f (ha)
- xf_f = area set aside by farm f (ha)
- q_{df} = binary variable to decide whether farm f will plant nonfood crop d
=1 if plant; =0 otherwise

Government Model(Leader)

$$\min_{\tau,b} \sum_{b \in B} \sum_{d \in D} \sum_{f \in F} \alpha_{db} r_{db} x n_{df} \tau_b - \gamma \sum_{f \in F} \sum_{d \in D} \eta_d x n_{df} \quad (1)$$

$$\text{subject to} \quad \sum_{f \in F} \sum_{d \in D} x n_{df} \geq \rho \theta \sum_{f \in F} w_f \quad (2)$$

$$\sum_{f \in F} \sum_{d \in D(b)} \alpha_{db} r_{df} x n_{df} \leq u_b \quad \forall b \in B \quad (3)$$

$$p_d = \max((\tau_b + v_b - \beta_{db} - \pi_b + o_{db})\alpha_{db}, 0) \quad \forall b \in B(d), d \in D \quad (4)$$

$$p_d \geq 0, \tau_b \geq 0 \quad \forall b \in B, d \in D \quad (5)$$

The objective (1) is to minimize the total value of tax credits provided by the government after the deduction of the exemption of γ to grow certain nonfood crops. Inequality (2) ensures the minimum percentage of the total arable land to be used for food crops. Constraints (3) limit the production of biofuel b to no more than L_b . These constraints explain the need for a certain amount of biofuel for the government, but ensure that biofuels that may have storage issues, such as ethanol, do not get produced excessively. The next constraint represents the industry's problem to price the nonfood crops appropriately. The nonnegative price for each nonfood crop (4) that the industry pay to the farms ensures itself get the expected profits π_b from the difference between the marginal profits, which include either from tax credit τ_b or selling price from biofuel v_b and from byproduct o_{db} , and the conversion costs β_{db} . The decision variable or the price coefficients should be nonnegative, specified by (5).

Farms Model(Follower)

$$\max \sum_{f \in F} \sum_{c \in C} m_{cf} x_{cf} + \sum_{f \in F} \sum_{d \in D} (p_d r_{df} + s_d - c_{df}) x n_{df} + \gamma \sum_{f \in F} x f_f - \sum_{f \in F} \sum_{d \in D} q_{df} t_{df} \quad (6)$$

$$\text{subject to} \quad \sum_{c \in C} x_{cf} + \sum_{d \in D} x n_{df} + x f_f \leq w_f \quad \forall f \in F \quad (7)$$

$$\sum_{d \in D} x n_{df} + x f_f = \theta w_f \quad \forall f \in F \quad (8)$$

$$x_{df} + x n_{df} \leq \kappa_{df} w_f \quad \forall f \in F, d \in D \quad (9)$$

$$x n_{df} \leq w_f q_{df} \quad \forall f \in F, d \in D, \quad (10)$$

$$x f_f \leq \delta w_f \quad \forall f \in F, \quad (11)$$

$$x_{cf} \geq 0, x n_{df} \geq 0, x f_f \geq 0, q_{df} \in \{0, 1\} \quad \forall c \in C, d \in D, f \in F \quad (12)$$

Function (6) is the objective function of the agriculture sector to maximize its profits by assigning the total arable land among food crops, nonfood crops and fallow lands. Inequality (7) confines the assignments to the total available land. Among them, certain amount of land is only used for nonfood crops and for keeping fallow in (8). Agronomic considerations are reflected in (9). For example, previous year's production levels, soil nutriment levels,

or the amount of labor and equipment available can limit the amount of land available for a type of crop. As such, for each nonfood crop, an upper bound may also be given as a percentage of available land on each farm. (10) is used to implement capacity constraints, where the value of q_{df} determines whether or not land is available at farm f to grow crop d . (11) gives the upper bound on the amount of land left fallow. (12) ensures that all the continuous variables are nonnegative and q_{df} are binary variables. The inclusion of fixed costs affects the optimization methodology and the complexity.

NUMERICAL EXAMPLE AND CONCLUSION

In this section, we will show a numerical example to explain the influence of having fixed costs on the modeling and optimization. Three farms of different size need to allocate their lands to plant 7 food crops. Some of them could be used as nonfood crops to produce biofuels. Wheat, corn and sugar beet could be transferred to ethanol by industry. Rapeseed and sunflower can produce another biofuel—ester. We used data consistent with [3] and filled any gaps in the data with further research into agricultural conditions in Europe. Given $\gamma = 193.2$, $\theta = 0.35$, $u_b = 100000$ and $\rho = 0.2$ and the rest are shown in TABLE 1.

TABLE 1: PARAMETERS

	π	v			α	β	o	r_{F1}	r_{F2}	r_{F3}	η
ester	60	72	ethanol	wheat	3.5	20.7	83	7.56	7.13	7	0
ethanol	120	60		corn	3.8	20.7	90	8	8	9	0
				sugar	1	13	0	7.5	7.43	7.4	1
F1	3000		ester	rapeseed	4.5	168	120	2.5	2.5	2.5	0
F2	14600			sunflower	4.7	168	105	2.29	2	2.4	0
F3	6300										

	m_{F1}	m_{F2}	m_{F3}	t_{F1}	t_{F2}	t_{F3}	c_{F1}	c_{F2}	c_{F3}	s
wheat	609.04	608.04	607.04	25600	29900	21200	756.7	799.5	812	37.47
barley	121.2	123.2	125.2							
corn	713	708	703	20000	20000	20000	200	200	200	35.75
sugar	200	200	200	20000	20000	20000	200	200	200	35.75
rapeseed	472.8	477.8	482.8	80000	20000	60000	200	200	200	37.47
sunflower	479.8	481.8	480.8	27300	27000	29000	273.93	270	290	37.48
peas	234.3	231.3	230.3							

We use grid search to solve the problem. The basic idea is the following:

- Assume that $\tau_{ester}, \tau_{ethanol} \in [0, 400]$. Starting with a step size of 25, we substitute the pair of τ into the equality constraints (4) to obtain appropriate values for p . Then we solve the lower level problem using the mixed-integer solver MINLP [11] to get the optimal assignment and calculate the corresponding government cost with (1) and the value of constraints (2)(3). Next, sort all the combinations, which have feasible (2)(3), in the ascending order of government cost. Choose the values of τ_{ester} and $\tau_{ethanol}$ that give the lowest government cost in each group of ten.

- For each pair of values $(\tau_{ester}^{(1)}, \tau_{ethanol}^{(1)})$ chosen in the previous step, redefine the search scale as $\tau_{ester} \in [\tau_{ester}^{(1)} - 25, \tau_{ester}^{(1)} + 25]$, $\tau_{ethanol} \in [\tau_{ethanol}^{(1)} - 25, \tau_{ethanol}^{(1)} + 25]$. Repeat the same procedure given above with a new step size of 5 and get the best point in each group of ten feasible pairs.
- Given each pair of values $(\tau_{ester}^{(2)}, \tau_{ethanol}^{(2)})$ chosen in the previous step, repeat the above procedure with a step size of 1 in the range $\tau_{ester} \in [\tau_{ester}^{(2)} - 5, \tau_{ester}^{(2)} + 5]$, $\tau_{ethanol} \in [\tau_{ethanol}^{(2)} - 5, \tau_{ethanol}^{(2)} + 5]$. Then we get a local minimum for the government cost.

TABLE 2: RESULTS OF THE NUMERICAL EXAMPLE

without fixed cost				with fixed cost			
x	F1	F2	F3	x	F1	F2	F3
wheat	0	1314	441	wheat	66	1314	441
barley	0	0	0	barley	0	0	0
corn	0	0	0	corn	0	0	0
sugar	0	0	0	sugar	0	0	0
rapeseed	66	0	0	rapeseed	0	0	0
sunflower	330	1314	756	sunflower	330	1314	756
peas	1749	6262	2898	peas	1749	6862	2898
xn	F1	F2	F3	xn	F1	F2	F3
wheat	231	146	378	wheat	165	146	0
corn	396	1606	504	corn	396	1606	504
sugar	0	0	0	sugar	0	0	0
rapeseed	198	1898	693	rapeseed	264	1898	693
sunflower	0	0	0	sunflower	0	0	0
xf	330	1460	630	xf	330	1460	630
τ_{ester}	125	$\tau_{ethanol}$	0	τ_{ester}	142	$\tau_{ethanol}$	0
gov_{obj}	3922031	$farm_{obj}$	9514154	gov_{obj}	4560863	$farm_{obj}$	9756897

Using the above procedure, we get the optimal land assignment for each food crop and each nonfood crop in each farm. We compare the results with fixed cost to those without fixed cost in the TABLE 2. When fixed costs are considered, the government should increase the tax exemption for ethanol from 125 to 142, leading to a 16.3% increase in the total cost for the government. Although the farms are faced with an increased number of costs, we observe that they do not produce less of the nonfood crops, but instead change their allocation in response to the fixed costs. As a result, the total profits of the farms increase by 2.55%. It seems that fixed costs influence the expenses of the government more than the profits of the farms, which is consistent with our intuition that the farmers could adjust their strategies easily according to the change of the cost structure while the government might take more efforts to react to the signal from the farmers. From the numerical example, we can see there is an significant impact of the fixed cost on the final decision and the profits which farms and the government share. We solved this problem using a simple grid search in this paper, and the goal of our future work is the development of more efficient solution methods.

Acknowledgements and Notes: We thank Bard, Plummer and Sourie for their assistance in locating the original model and data files for [3]. References available upon request from authors.

MODELING TRAVELERS' BEHAVIOR IN THE PRESENCE OF HOT LANES USING A TRAFFIC ASSIGNMENT METHOD

Soheil Sibdari

University of Massachusetts Dartmouth
285 Old Westport Rd., Dartmouth, MA 02747
ssibdari@umassd.edu
508-999-8019
(Corresponding Author)

Mansoureh Jeihani

Morgan State University
1700 E. Cold Spring Lane, Baltimore, MD 21251
mjeihan@eng.morgan.edu
443-885-1873

ABSTRACT

In this paper, we highlight the fact that using value pricing for HOT lanes, the congestion level of regular lanes can also be controlled. Travelers adjust their departure time not only based on the congestion levels but also based on the pricing pattern that is used for the HOT lane. A user equilibrium method is exploited to calculate the number of travelers in each route during a peak period. Using this method, we provide a numerical study to determine the distribution of travelers among a peak period under different pricing menus.

KEYWORDS

Congestion pricing, HOT-lane management, traffic assignment

1. INTRODUCTION

Traffic congestion is a serious problem in urban areas challenging transportation policy makers, and imposes an important part of the total cost of transportation. Recently, Transportation Demand Management is used to manage traffic congestion more efficiently. In this approach, travel demand is assumed to be a quantity, requested by road users, whose size and distribution can be controlled using different policies. Congestion pricing along with flexible work hours, ride-sharing, and public transportation incentives are the policies that can manage travel demand.

As a special case of congestion pricing, some researchers have addressed the concept of value pricing, which lets travelers to choose between a free but congested and a priced but free-flowing roadways (Small and Yan [1]). This pricing policy gives a chance to those single drivers to pay for more highly valued services. These express lanes that were introduced by Fielding and Klein [2] are also called High Occupancy Toll (HOT) lanes. The existing HOT lanes implemented in the U.S. are Interstate 15 in San Diego, California, Quick Ride System on the Katie Highway and the Northwest Freeway in Houston, Texas, and the SR 91 Express Lanes in Orange County, California (Evans, et al. [3]).

Value pricing can also resolve some other issues associated with High Occupancy Vehicles (HOV) lanes. For example, some studies demonstrate that in some corridors HOV-2 lanes are under-utilized while in some other corridors they are very congested, and in the case of congested HOV-2, changing HOV-2 to HOV-3 makes them almost empty lanes. Using HOT policy, low occupancy vehicles are allowed to enter by paying a toll while carpools or buses are allowed to use the lanes for free or at a discounted toll rate. Therefore, the congestion level of the under-utilized HOV-2 or HOV-3 lanes can be controlled by using them as HOT lanes and charging a right toll to use these lanes.

Different types of objective function have been utilized in the existing literature to develop value pricing models. The first type of models deals with second-best tolling policy to the express lane using which social welfare is maximized subject to the zero-toll constraint on the alternative roadways. The optimal uniform second-best tolling policy is set to be a weighted average of the marginal external congestion costs between non-carpooling and carpooling vehicles. The second type of criteria deals with "profit maximizing" tolling policy which maximizes revenue subject to the zero-toll constraint on the other roads. Small and Yan [1] compare the outcomes of the profit-maximization policy with the second-best pricing policy and find out that the profit-maximizing policy sets the toll higher, and social benefit lower, than the second-best pricing policy. The third type of pricing regime sets the toll high enough to keep the flowing of the priced roadway at a minimum specified speed. This policy achieves social optimum like the second-best pricing policy, but with an additional constraint designed to guarantee a minimum level of service in the express lane. Small and Yan [1] model this pricing policy and find out that the travelers' behavior under this policy is almost the same as

revenue-maximizing policy. For detail review of pricing policies, we refer the readers to Yang and Huang [4], Braid [5], and Verhoref et al. [6].

Unlike the current literature, we are interested in the impact of HOT lane pricing on the congestion level of regular lane during a peak period. Consider a pricing policy that charges higher tolls for the use of HOT lanes during the peak period and lower tolls for the shoulders of the peak period, where the traffic is relatively low. In this situation, some travelers who plan to depart during the peak period may change their departure time and use the HOT lane in periods with a lower toll while enjoying lower traffic levels. For example, consider a regular lane and a HOT lane connecting two points with the same length and free-flow travel time. The toll rates for the use of HOT lane are \$3 from 4:00pm till 5:00pm and \$6 from 5:00pm till 6:00pm. A traveler who initially plans to depart at 5:00pm and is not willing to pay \$6, might change her travel time to 4:00pm in order to use the HOT lane by paying \$3, rather than using the regular lanes at 5:00pm.

Overall, using a right pricing policy, more travelers have incentives to change their departure time from congested periods of the peak period to the shoulders of the peak period. This behavior smoothes out the departure rate during the peak period and lowers the traffic level of regular lanes. Addressing this observation is the focus of this study. We employ a user equilibrium method to compare the impact of different well-known pricing policies on the travelers' behavior. We also provide pricing examples which control the traffic of regular lanes better than the existing models.

We begin with a brief review of the literature on value pricing in Section 2. Section 3 describes the problem, presents our assumptions, and provides the solution method. In Section 4, we employ a user equilibrium technique to determine the distribution of travelers during the peak period under different pricing policies. Conclusion remarks are listed in Section 5.

2. LITERATURE REVIEW

Value pricing is well studied in the literature in two types of static and dynamic pricing. Static pricing assumes that travel demand and costs are not time sensitive. Therefore, the inter-temporal impact of the present toll on the long run congestion level is not considered. Most of studies in this area employ the theory of marginal cost pricing to estimate an optimal congestion toll on transportation networks. This pricing policy charges higher fees on road users who travel during peak periods. The road users should pay a toll equal to difference between marginal social costs and marginal private costs. Dafermos and Sparrow [7] determine a system of optimal tolls on general transportation networks using a static congestion pricing. Yang and Meng [8] calculate static congestion pricing using marginal social cost and marginal private cost of a travel, where roads are modeled as bottlenecks.

In dynamic pricing methods, the travel demand, travel costs, and the toll level vary over time. In a seminal paper by Vickrey [9] a dynamic congestion pricing is

modeled by considering a single bottleneck with fixed capacity and fixed number of travelers. He considers work-related trips and assumes that the departure times of all travelers are the same, and each driver faces a trade-off in what time to go to work. If the traveler leaves early, she faces no congestion but she arrives soon, which occurs a cost due to arriving early, and if she leaves such that she arrives on time, she might face congestion and face the cost of arriving at work inconveniently late. Using this idea, Vickrey [9] obtains the equilibrium of the length of the queue over time when no driver can reduce her trip price by changing her departure time.

Arnott et al. [10] extend Vickrey's model by considering heterogeneous travelers. They develop a deterministic mathematical model to determine the impact of optimal time-varying tolls on social welfare. They also examine different tolling schemes such as uniform and step-function on the system efficiency. Using a time-dependent congestion toll and an optimal congestion pricing, they achieve a more uniform departure rate and therefore reduce both congestion and traveling time for the commuters.

Another approach to determine dynamic tolls deals with both time and space dimensions of road use. Arnott et al. [11] consider a fixed demand in a network with parallel routes and use dynamic traffic assignment to examine the impact of different pricing regimes on the congestion level of the network. Cary and Srinivasan [12] address system marginal costs, user externality costs, and optimal congestion toll using a non-linear programming approach. Yang and Huang [4] address a time-varying pricing model of a road bottleneck when demand is elastic and deterministic. They employ the continuous-time optimal control approach to solve this problem. They also determine the optimal use of bottleneck by maximizing social benefit over the time horizon of study.

Liu and McDonald [13] use economic theory and simulation models to compare first-best, second-best, and no-toll solutions for a model with two routes and two time periods (peak period and pre-peak period). First, they found that the second-best pricing policies are effective in reallocating traffic volumes but less effective than the first-best tolling policy. Second, the amount of second-best is less than first-best tolling policy. Third, the social welfare obtained from the second-best tolling policy is smaller than the welfare gained by first-best tolls.

3. MODEL

There are two roadways connecting points A and B with the same length L and the same free-flow travel time. One of these roadways is toll-free while the other one is subject to a toll collection, which we call them roadways NT and T , respectively. The only points of access and egress for both roadways are A and B and travelers can only choose between these two routes to go from A to B . The high occupancy vehicles (HOV) can use both roadways and single occupancy vehicles (SOV) can use the NT roadway for free and the T roadway by paying a fee, which can vary over time.

In this study, we consider a total of N travelers in single occupancy vehicles (SOV) with home-based-work (HBW) trips who plan to travel from point A to point B during a peak period. HBW trips are the most important trip types during the peak period because they are less flexible in departure time. Each person should be at work at a certain time of day. Besides, these types of trips are usually made by SOVs. The length of the peak period is n , which is divided into n time slots with unit lengths. Without loss of generality, we consider an odd number of periods and name the middle time slot as time slot 0 and other time slots accordingly. For example, a time period with $n = 5$ from 3:00pm till 8:00pm can be divided into 5 time slots 3:00-4:00, 4:00-5:00, 5:00-6:00, 6:00-7:00, and 7:00-8:00 that we name these time slots as -2, -1, 0, 1, and 2 respectively.

Each traveler has a preferred time slot, and since we only consider HBW trips, all travelers prefer a specific time slot, say time slot 0. However, depending on the congestion level and the HOT lane toll price, travelers might change their departure time that causes them a disutility. Since we only consider a single type of travelers, the magnitude of this disutility is fixed among travelers, but it varies by departure time. For example, departing in time slot 0 imposes no disutility while traveling in other time slots imposes positive disutility, with the disutility of traveling in time slots 2 and -2 be greater than the disutility of traveling in time slots 1 and -1. The disutility of traveling in time slot -1 is equal to the disutility of traveling in time slot 1, similarly for time slots -2 and 2. In addition to the disutility, travel time and toll level are other parameters of travelers' cost. Travelers choose a roadway and a time slot in order to minimize their total travel cost. We now use a traffic assignment model to find an equilibrium solution specifying number of travelers on each roadway in each time slot.

Traffic assignment models are employed to assign the flow of traffic between an origin and a destination among different routes. A wide variety of traffic assignment models have been applied in the literature. System optimal and user equilibrium are among the most popular models. System optimal models find an assignment that minimizes total network travel time. Using these models, no traveler can change her route without increasing total system travel time. In user equilibrium approach no traveler can improve her travel time by switching routes. We employ the user equilibrium approach in our study.

To compute user equilibrium, we use Frank-Wolfe convex combination algorithm (Frank and Wolfe [14]), which is widely used in the literature. It assigns all the flows to the initial shortest path, and then iteratively updates the link costs using a volume delay function, where volume delay function presents the relationship between the cost of traversing a link and the flow of the link. It then finds a new shortest path between each origin and destination, and assigns a convex combination of flow to the new and old shortest paths. To apply the algorithm, we first need to define routes and their generalized cost function.

Consider each roadway (T or NT) at each time slot as a separate link with the same length. Thus, the network contains $2n$ links with the same length and free-flow speed. For instance, time slot -2 of the regular lane is one link and time slot -1 of the

regular lane is another link and so on. The volume delay function for each link is defined as follows.

$$C_i^j(V_i^j) = \delta p_i^j + \gamma D_i + \varphi t_i^j [1 + \alpha_i^j (\frac{V_i^j}{Cap_i^j})^{\beta_i^j}] \quad \text{for } i = -(n-1)/2, \dots, 0, \dots, (n-1)/2 \text{ and } j =$$

$T, NT.$

where:

C_i^j : generalized cost of traveling on the link presenting time slot i and roadway j

D_i : cost associated with the disutility of choosing the link presenting time slot i rather than time slot 0

t_i^j : free-flow travel time of the link presenting time slot i and roadway j (in minutes)

p_i : toll price during time slot i for the links presenting toll road

V_i^j : flow of the link presenting time slot i and roadway j

Cap_i^j : capacity of the link, which presents time slot i and roadway j

$\alpha, \beta, \delta, \gamma, \varphi$: coefficients

t_i^j is calculated as follows.

$$t_i^j = \frac{L_i^j}{FFS_i^j} * 60$$

L_i^j : length of the link presenting time slot i and roadway j (in miles)

FFS_i^j : free-flow speed of the link presenting time slot i and roadway j (in miles per hour, mph)

The above volume delay function is called the Bureau of Public Roads (BPR) function with generalized cost. BPR is one of the most commonly used volume delay functions in which the cost of traversing a link is simply the link travel time which is calculated as a function of free-flow travel time of a link and the ratio of link flow to link capacity. However, in BPR with generalized cost, it is assumed that traversing a link conveys other costs besides link travel time. We assume these costs are the possible toll and the disutility of selecting a time slot other than time slot 0. Now, the stepwise explanation of the applied Frank-Wolfe algorithm can be explained using the defined notations as follows.

1. Initialization: perform all-or-nothing assignment based on $C_i^j(0)$. This gives flow vector $V_i^j(1)$. Set the iteration number m to 1.
2. Update Travel Cost: update the link travel cost $C_i^j(m) = C_i^j(V_i^j(m))$
3. Direction finding: perform all-or-nothing assignment with the updated $C_i^j(m)$. This gives the auxiliary flow vector $Y_i^j(m)$

4. Line search: find λ that solves
$$\min_{0 \leq \lambda \leq 1} \sum_j^{V_i^j(m) + \lambda(Y_i^j(m) - V_i^j(m))} \int_0^{C_i^j(u)} du$$
5. Move: set the flows to $V_i^j(m+1) = V_i^j(m) + \lambda(Y_i^j(m) - V_i^j(m))$
6. Convergence test: if $\lambda < \varepsilon$, stop; otherwise go to Step2, where ε is a very small number, e.g. 0.0001.

The above algorithm yields an equilibrium solution in which travelers are distributed among the $2n$ links with minimized possible cost. In addition, in this equilibrium solution, all links have the same cost. Since each link presents a time slot and a roadway, the N travelers are distributed among each time slot on either toll-free or tolled lane such that they experience the same travel cost including travel time, the disutility of choosing the slot, and the toll.

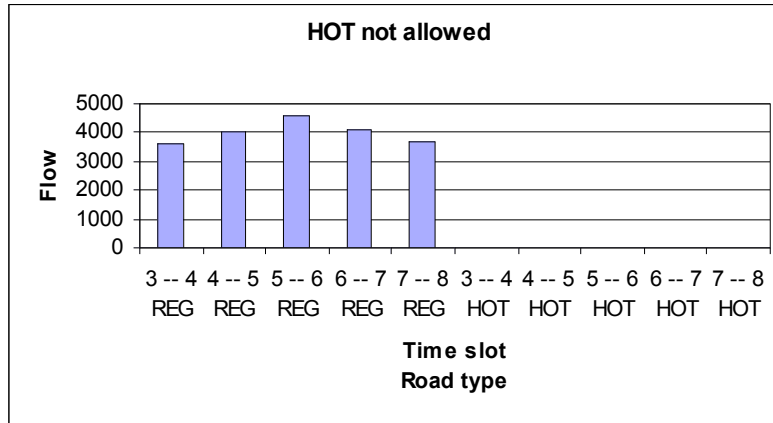
4. NUMERICAL STUDY

In this section we use a numerical study to illustrate our results. We consider a total of $N = 20,000$ travelers who plan to travel from an origin to a destination with two alternatives of regular lane and HOT lane. We consider $n = 5$ from 3:00pm till 8:00pm and we divide this time period into 5 equal time slots of 3:00-4:00, 4:00-5:00, 5:00-6:00, 6:00-7:00, and 7:00-8:00 and we call them as time slots -2, -1, 0, 1, 2, respectively. As explained in Section 3, for our user equilibrium model, we name each time slot of each lane as a link with the same length ($L_i^j = 10$ miles) and the same capacity ($Cap_i^j = 2800$ vehicles per hour per lane), and the same free-flow speed ($FFS_i^j = 65$ mph), for all i and j , $i = -2, \dots, 2$ and $j = T$ and NT . Each traveler has a total of 10 choices of routes. We assume $\gamma = 1$ and $D_i = 20, 10, 0, 10, 20$ as the disutilities of changing to time slots i , for $i = -2, \dots, 2$. In the following, we first determine the distribution of travelers among time slots of regular lane in the case that the HOT lane is not available and then we compare the impact of different pricing policies on the travelers' choice of route and departure time. Then, we study the distribution of travelers in the presence of a HOT lane.

4.1 HOT Lane Not Available

In this case, we assume that HOT lane is not allowed and that SOVs cannot use the express lane. Therefore, travelers have five choices of time slots of the regular lane. The result of this case provides us the initial distribution of travelers among available time slots. Figure 1 presents the flows on each time slot of regular lane, and as expected, the distribution of travelers among time slots has a bell shape. In equilibrium, the generalized costs of each time slot are the same and that no traveler has incentive to switch to another time slot. Because there are only two factors of disutility and travel time involved in calculating the generalized cost function, travelers face a trade-off between changing the time slot and having more travel time.

FIGURE 1 The Flows in Regular Lane When HOT Lane Is Not Allowed



4.2 Static Toll Pricing Models

For the rest of our numerical analysis, we let SOVs to use the HOT lane. In this section we use static pricing models, which charge a fixed rate throughout the peak period. To illustrate the impact of toll level on the travelers' decision, we first study the case of free HOT lane and then we study the case with a fixed toll.

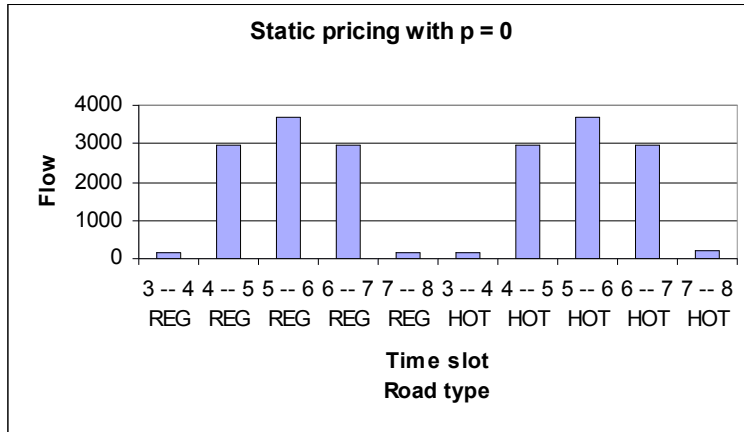
4.2.1 Free HOT Lane

When SOVs can use the HOT lane for free, it can be treated as if a new facility is constructed as a regular roadway. Figure 2 presents the distribution of flows among the ten routes. As expected, travelers are distributed evenly between two roadways and they mostly choose time slots 0, 1, and -1, since there is no toll and they prefer not to change their original departure time. However, few travelers choose time slots 2 and -2 and enjoy free-flow travel time. Based on Highway Capacity Manual, level of service (LOS) of slots -2, -1, 0, 1, and 2 are A, C, D, C, and A respectively, where Level of Service is a measure to determine the quality of service of transportation facilities, grading alphabetically from A to F with A being the best and F being the worst.

4.2.2 Fixed Toll HOT Lane

When we charge a fixed toll of \$2, as demonstrated in Figure 3, some travelers change their departure time from the middle time slots to the shoulders. As a comparison with case 4.2.1, the flows on regular lane in time slot -1, 0, and 1 remains almost the same, with flows of 2961, 3712, and 2985 when $p = 0$, and flows of 3057, 3766, and 3060 when $p = \$2$. On the other hand, the number of travelers using the HOT lane in time slots -1, 0, and 1 drops significantly, with flows of 2966, 3709, and 2963 when $p = 0$, and flows of 1729, 3051, and 1715 when $p = \$2$.

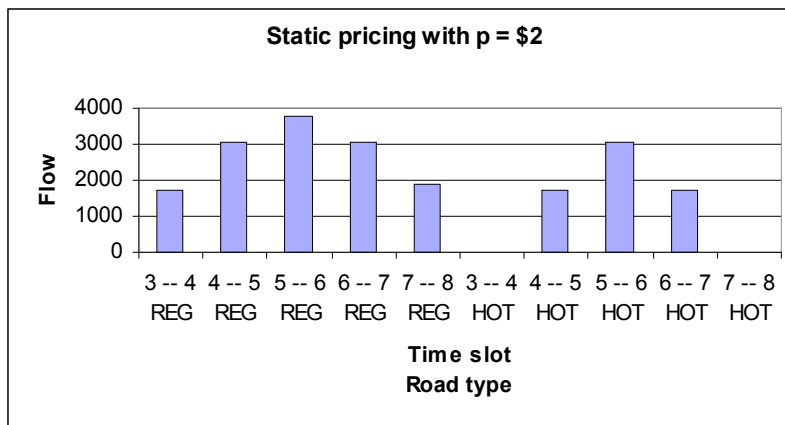
FIGURE 2 Route Flows When HOT Lane Is Free



By charging a fixed toll of \$2, some travelers change their departure time from time slot -1, 0, 1 and leave earlier or later in order to pay no toll while enjoying low traffic with LOS = B in time slots -2 and 2 of the regular lane. The travelers have no incentive to choose time slot -2 and 2 of the HOT lane and pay \$2 while the same time slots have low traffic in the regular lane for free.

As a comparison with Section 4.1, by adding HOT lane, many travelers change their departure time and use the HOT lane by paying a toll and traveling during their preferred departure time. More travelers switch from the shoulder time slots of the regular lane to the preferred time slots of HOT lane.

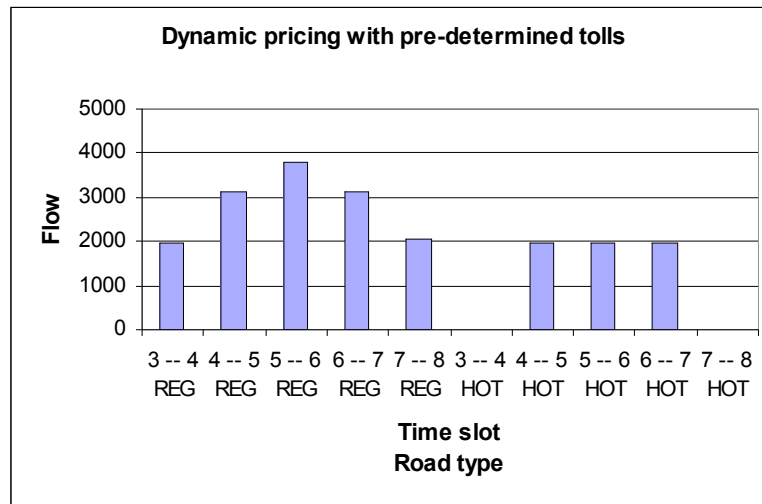
FIGURE 3 Route Flows When Toll is Fixed at \$2



4.3 Dynamic Toll Pricing Model

In this section, the customers' behavior is studied when the toll levels vary over the peak period. We first use a pre-determined step function that is used for most of the current HOT lanes in the nation. Then we use a dynamic pricing model based on level of service to calculate the toll for each time slot. Using our proposed pricing menu, more

FIGURE 4 Route Flows When Toll Changes in Each Time Slot



travelers change their departure time and that the congestion level is more uniform during the peak period compare to the congestion level under other pricing policies.

4.3.1 Pre-determined Dynamic Pricing

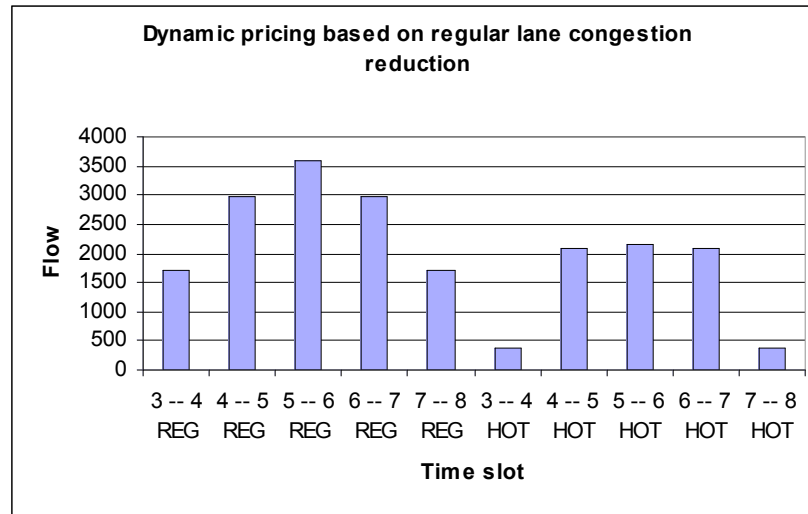
Using the pricing policy that is used in I-15 (I-15 FasTrak Program, 2007), we fix the toll levels of time slots -2, -1, 0, 1, and 2 at \$1, \$2, \$4, \$2, and \$1, respectively. The results of the numerical analysis are summarized in Figure 4. As we can see, the total flows on the HOT lane during the time period decreased by 581 compare to the total flows on HOT lane when a fixed toll of \$2 is charged. This is because a higher toll of \$4 is charged for time slot 0 and more of travelers depart in time slot 0 compare to time slots -2 and 2 . In addition, because of the higher toll of \$4 on time slot 0, some travelers who initially plan to travel in time slot 0 switch their departure time to time slots -1 and 1 of HOT lane. Some other travelers switch to regular lane and change their departure time toward the shoulder of the peak period.

4.3.2 Proposed Dynamic Pricing

In this section, we use a pricing policy under which the congestion level of the regular lane is decreased while the HOT lane remains at level of service C. Since, our objective does not contain profit maximization, the toll levels might be lower than the pre-determined toll levels described in Section 4.3.1. Figure 5 shows the resulting flows for the toll policy 0.5, 1.5, 3, 1.5, and 0.5 for time slots -2, -1, 0, 1, and 2, respectively.

Comparing the results with Section 4.3.1, first, the total number of travelers using the regular lane is decreased by 1146. Second, more travelers change their departure time from middle time slots to time slots -2 and 2 of the HOT lane. The number of travelers changing their departure time from time slot -1, 0, and 1 of the regular lane are 174, 327, and 170. These travelers depart in time slots -2 and 2 of the HOT lane with 365 and 366 travelers currently depart in these two time slot. This result clearly illustrates our conclusion that by charging a right pricing menu for the HOT lane during the pea

FIGURE 5 Route Flows When Toll Changes Dynamically Based on the Reduction of Regular Lane Congestion



period, the congestion level of the regular lane can be reduced while the traffic level of the HOT lane remains in an acceptable level.

5. CONCLUSIONS

In this paper travelers' behavior is studied in a multilane highway in which one lane is tolled, HOT lane, and the other lanes are free. We investigate the impact of HOT lane pricing on the congestion level of regular lanes during a peak period. We demonstrate that the HOT lane prices make some travelers to change their departure time from peak period to the shoulders of the peak period and use the HOT lane instead of the regular lane. A user equilibrium method is employed to determine the distribution of travelers during the peak period between HOT and regular lanes. Using this method, we compare the impact of different well-known pricing policies on the behavior of travelers. We also provide a pricing menu that controls the traffic of regular lane better than the existing pricing methods in the literature.

As an extension to this study, one can develop an optimization model to calculate a pricing menu using which the of travelers' departure rates are more uniform during the peak period. Another extension of this study is to develop a multi-objective optimization model that both reduces the regular lane congestion level and maximizes the total expected profit. Note that a pricing method based on social welfare maximization does not completely address our problem, since it aggregates the total utility for all travelers and the social planners, which does not deal with the distribution of travelers during peak period. Finally, one can define multiple classes of travelers who face different value of time and also different disutility levels of changing departure time.

REFERENCE

- [1] Small, K.A. and J. Yan. The Value of Value Pricing of Roads: Second-Best Pricing and Product Differentiation. *Journal of Urban Economics*, Vol. 49, 2001, pp. 310-336.
- [2] Fielding, G. J. and D. B. B. Klein. High Occupancy/Toll Lanes: Phasing in Congestion Pricing a Lane at a Time. *Reason Foundation, Policy Study*, Vol. 170, 1993.
- [3] Evans, J., K. Bhatt, and K. Turnbull. *TCRP Report 95*, Chapter 14-Road Value Pricing, 82nd *TRB Annual Meeting*, Washington, D.C, 1993.
- [4] Yang, H. and H. Huang. Carpooling and Congestion Pricing in a Multilane Highway with High-Occupancy-Vehicle Lanes. *Transportation Research A*, Vol. 33, 1999, pp. 139-155.
- [5] Braid R. Peak-Load Pricing of a Transportation Route with an Unpriced Substitute. *Journal of Urban Economics*, Vol. 40, 1996, pp. 179-197.
- [6] Verhoef, E., P. Nijkamp, and P. Rietveld. Second-Best Congestion Pricing: the Case of an Untolled Alternative. *Journal of Urban Economics*, Vol. 40, 1996, pp. 279-302.
- [7] Dafermos S. and F. Sparrow. Optimal Resource Allocation and Toll Patterns in User-Optimized Transportation Network. *Journal of Transportation Economics and Policy*, Vol. 5, 1971, pp. 198-200.
- [8] Yang H. and Q. Meng. Departure Time, Route Choice and Congestion Toll in a Queuing Network with Elastic Demand. *Hong Kong university of science and technology, working paper*.
- [9] Vickrey W. Congestion Theory and Transportation Investment. *American Economic Review*, Vol. 59, 1969, pp. 251-261.
- [10] Arnott R. A. de Palma, and R. Lindsey. A Structural Model of Peak-period Congestion: A Traffic Bottleneck with Elastic Demand. *American Economic Review*, Vol. 83, 1993, pp.161-179.
- [11] Arnott, R., A. dePalma A., and R. Lindsey. Departure Time and Route Choice for the Morning Commute. *Transportation Research B*, Vol. 24, 1990, pp. 209-228.
- [12] Carey, M. and A. Srinivasan. Externalities, Average and Marginal Costs, and Tolls on Congested Network with Time-varying Flows. *Operations Research*, Vol. 41, 1993, pp. 217-231.

[13] Liu, L. N., and J. F. McDonald. Economic Efficiency of Second-best Congestion Pricing Schemes in Urban Highway Systems. *Transportation Research B*, Vol. 33, 1999, pp.157-188.

[14] Frank, M. and P. Wolfe. An algorithm for quadratic programming. *Naval Research Logistics Quarterly*, Vol. 3, 1956, pp. 95-110.

INTERNET MARKETING AND CUSTOMER RELATIONSHIPS MANAGEMENT: A TWO STEP COMMUNICATIONS STRATEGY

Michael V. Laric, University of Baltimore mlaric@ubalt.edu
Peter M. Lynagh, University of Baltimore plynagh@ubalt.edu

Keywords: Internet Marketing, Search Engine Marketing, Customer Relationships management

ABSTRACT

This paper proposes a two step communications approach for Internet Marketing (IM) In the first step the emphasis is on attracting and capturing new customers using Websites and Search Engine Marketing (SEM). In the second step customers who have registered on the company Website have a “cookie” placed on their computer so in future visits the consumer would be admitted to the website “automatically.” This starts a Customer Relationship Management System (CRMS) which turns these customers into repeat customers by making it easy to come back. Amazon.com is a site that uses this approach in its marketing strategy.

INTRODUCTION AND OVERVIEW OF THE MODEL

The first step, in our proposed approach is comprised of SEM activities which incorporate two elements: organic search engine optimization (SEO) and the pay per click (PPC) advertising search. Once consumers find the sought after website, the business or other organization that built that website has an opportunity to move to step two by planting cookies which help identify the visitor to the website. The first step helps users locate information on topics of interest. The second helps consumers in becoming frequent visitors to the site.

In the first step, the search engines (such as Google and Yahoo!) search through the Internet, then store information about myriad websites (pages) which they have retrieved. The pages are analyzed and indexed using key words extracted from titles, headings, and special fields. When a user enters their search words or phrases, the search engine returns a listing of Web pages that best match the searchers input data.

SEO is the practice of using HTML codes, web page copy, site navigation, and more, in order to improve the site’s standing in the results returned by the search engines. PPC

advertising is another method used by search engine to attract the consumer to the advertised website.

The next step uses cookies which allow websites to store a small chunk of information on the customers' computer when they visit a specific website. After placing such a cookie on the customer's computer, a visit to the same site would let that site retrieve the cookie and recognizes the customer by the information supplied when he or she first registered with the site. That site can then customize the web pages the customer will see on the site in a way that would (hopefully) match their preferences.

In the second step, cookies are implanted into the users' computers. There are those who do not "surf" the Web seeking information but their numbers are dwindling and they are increasingly rare in today's high-tech world of business and IM. Surfers are usually seeking information, purchasing items or services, connecting with other surfers – socially or for business, etc. When they contact a Website, the server (driving the Website) frequently asks for some information.

Cookies have benefits and and customer relationship management system make life on the Internet easier for many users, and can be a time saver. A shopper can go to the alumni gift store of their alma mater and that entire shopping and shipping experience is fast and uncomplicated because of the information available to the store. In the world of B2B information made available by cookies can make the supply-chain function in a more efficient manner. Cookies can also maximize the value of a Web site. Information is available regarding what pages are looked at, how much time is spent on a particular page, and other pertinent information. Modifications can be made on the site improving the marketing experience for all parties.

If used properly, cookies can go a long way toward building long-run relationships between the firm and its customers. These strong relationships become a key factor in developing, nurturing and solidifying value for all those involved.

Is all advertising truthful? Are all sales persons honest? Are all sales promotions faultless? The answer to these, and other questions, is, of course, no. There will always be those individuals and firms operating in the darker side of marketing and IMC. The same can be said for those using cookies, i.e., some will collect data and then misuse that information.

Cookies have some disadvantages; they can be used to send blanket advertisements that annoy users. Spam will not only annoy the user it can lead to overloaded computers and possible breakdowns. There is also the danger that personal information

can be collected and used for nefarious purposes. It is also possible to steal cookies and use them improperly. This can happen in several ways. One is called packet sniffing where information may be read by computers on the network other than its receiver and the sender. Another problem is cross-site scripting where browsers send cookies to servers that should not receive them. A third problem involves the process of changing the value of cookies and this is called cookie poisoning. All these are very serious problems, which can lead to abuses and identity theft; however, there are procedures available to help prevent the occurrence of such problems.

Cookies alone will seldom if ever be enough to achieve Integrated Marketing Communications (IMC) objectives and help build brands and achieve long-term customer relationships. Internet marketing is an exciting part of promotion in the later stages of the first decade of the 21st century. However, other new and exciting options are available as well as clever and creative things taking place.

INTERNET MARKETING AND CUSTOMER RELATIONSHIPS MANAGEMENT

Figure 1 shows the Internet Marketing and Customer Relationships Management System Approach. There are three large circles corresponding to the main three components of IM. Three are customers, search engines and companies or organizations.

(A) **Customers** are Web using end-users or resellers and/or specific segments of customers (B) **Search engines** help customers look for information, ideas, products or services (C) **Companies or organizations** with Web-sites try to attract customers by using IM strategies

The three main components (A, B, and C) of IM have a micro and a macro environment. The micro environment is located within the three components and relates to the common (intersecting) areas, of which there are 4 more areas. These are:

(D) **Search Results and Loyalty** result when individuals use search engines to find information. Search engines sell advertising and or try to build traffic, they attract searchers by offering a variety of features designed to create loyalty. In that regard, search engine companies are very similar to other companies on the Internet. The intersection between (A) customers and (B) search engines is where the search engines are able to give consumers a wealth of useful information about ecommerce and other items of interest, and generate repeat searches by satisfied customers.

(E) Search Engine Optimization (SEO) and Pay Per Click (PPC) are ways of attracting customers. Companies and Web-site designers try to build attractive sites which will both attract customers, and improve the site's chances of coming up on the first few pages of customers' searches. Companies can bid for key words which customer search for; when the word is entered into a search, the highest bidder's Website shows up in the search and the company pays per click. The intersection between (B) search engine and (C) the company allows for these relationships.

(F) Information Exchange – the intersection of (A) customers and (C) the company allows the exchange of information, which can also lead to transactions, and which takes place directly between companies and customers

At the intersection of areas D, E, and F above is the fourth (and most important) intersection: (G) Customer Relationship Management System – It is the culmination of the interactions among all the elements and forms the center of the intersections of ABC through DEF

The area enclosed by the three main components (A, B, and C) of IM has an external macro environment, which consists of five elements. These can impact relationships among the IM main components. The five elements are:

1. Cultural Context – here the global cultural context impacts the forces and relationships within the IM environment. For example the Chinese Government has had its own regulations imposed on the Internet and on search engines in a way that is quite different from the U.S. environment.

2. Company strategy – there are major differences among Internet strategies as far as the philosophy of operations and what companies expect from their IM strategy. In the business to business (B2B) environment the purpose of the company's Web-site may be quite different from the one used for business to consumer (B2C) environment. Regardless of the environment, some companies operate both virtual and non-virtual operations, as “click and Mortar” businesses. There are those who rely on others to offer the physical aspects and operate as virtual, or “click” only. Amazon.com is a virtual site with no physical outlets.

3. Competition – the competitive arena in the IM world changes dramatically and rapidly, as new industries emerge practically overnight. The search engine corporation was virtually unknown just a short decade ago, and in less than a decade its major market share leader – Google – has a capitalization ranking in the top ten of U.S. corporations.

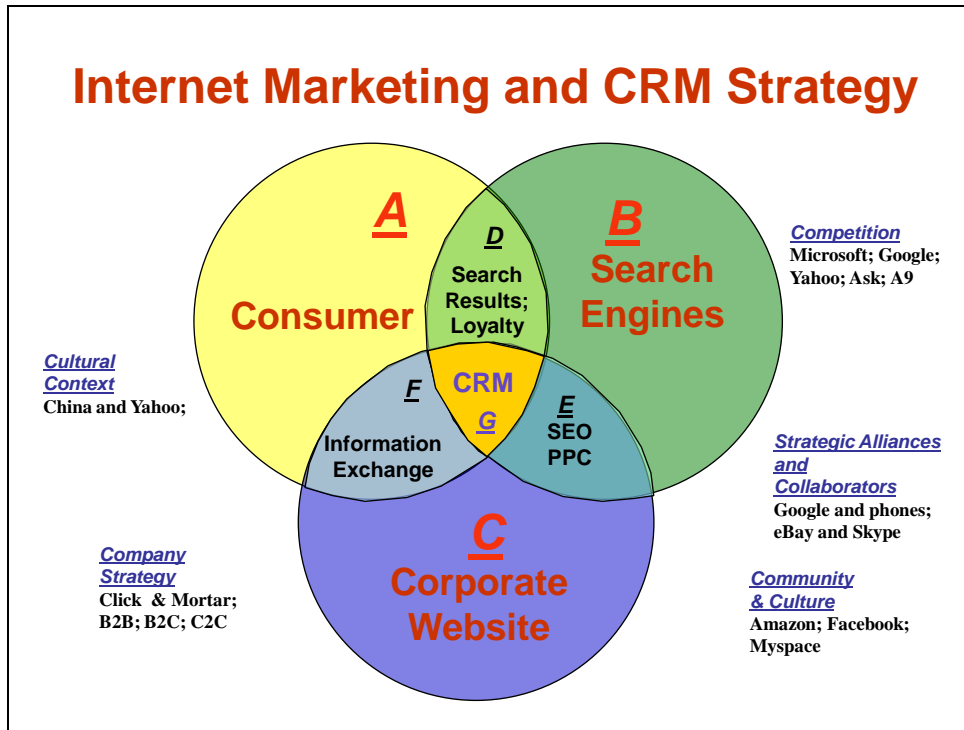


Figure 1 Internet Marketing and Customer Relationships Management

Regardless of the environment, some companies operate both virtual and non-virtual operations, as “click and Mortar” businesses. There are those who rely on others to offer the physical aspects and operate as virtual, or “click” only. Amazon.com is a virtual site with no physical outlets.

4. Competition – the competitive arena in the IM world changes dramatically and rapidly, as new industries emerge practically overnight. The search engine corporation was virtually unknown just a short decade ago, and in less than a decade its major market share leader – Google – has a capitalization ranking in the top ten of U.S. corporations. While competition between Google and Yahoo! was traditionally accepted as a battle for search engine market hegemony, more recently however Google entered the battle for cell phone services, pitting Google against AT&T and other cellular phone services. Likewise Google entered the battle for Web-based software applications, pitting Google against Microsoft.

5. Collaboration and Strategic Alliances – the ability to form relationships across industry boundaries has affected the competitive scene. Google sought the collaboration of phone manufacturers when it developed the new software for cellular phones, due out in mid-2008. eBay purchased Skype, the Internet telephony company and Microsoft invested in Facebook while Google cut a deal with Myspace.com.

6. Community and Culture – creating the community of users is now a recognized strategy considered essential to a variety of endeavors. Book reviews by users are now an important aspect of Amazon.com making everyone personally invested in “building” the site. Other “community” driven sites include Facebook.com, YouTube.com, MySpace.com, Friendster.com and others. Travelocity also carries reviews of hotels and travel destinations as part of its “unbiased” “reader generated” contents.

THE TWO STEPS AND CRM

The search engine marketing aspects of Internet marketing and customer relationships management approach moves through step one when the customer initiates a contact. (see figure 2.) An arrow from the customer (SEM) to any of the circles depicts that move. The approach moves to step two when the IM strategy establishes a connection using a cookie. This can be depicted by an arrow originating at any of the seven areas in Figure 1, to the customer. CRM would be depicted by both arrows, or by having an arrow pointing in both directions. Building a brand and developing long term mutually beneficial relationships require the blending of many IMC media. Two of the relatively new methods closely associated with Internet marketing would be Mobile Marketing (M Marketing) and word-of-mouth. Other forms of IM leading to this CRM model include: **Viral Marketing** is defined as *the Internet version of word-of-mouth marketing---Web sites, e-mail messages, or other marketing events that are so infectious that customers will want to pass them along to friends.* (Kotler 2008) Burger King’s “Subservient Chicken” campaign is an example. (www.subservientchicken.com;) other examples are **Social Networks**; Facebook.com, Friendster.com, LinkedIn.com are three examples of sites which link individuals and allow them to share personal or professional information with each other. Facebook.com is experimenting with “Beam” which uses individual’s shopping information, then shares the information with the purchaser’s friends. Another is **Mobile Marketing** which uses cellular phones and PDA’s for “...content delivery and direct-response within a cross-media marketing communications program. (Liesse, 2007, 6). With over 3 billion cell phones in a world of 6 billion people this CRM vehicle is likely to grow in importance

REFERENCES

Kotler, Phillip, and Gary, Armstrong (2008), Principles of Marketing, 12th (ed.)

Liesse, Julie, (2007), “Mobil Moves Forward” Special Advertising Section, Advertising Age (November) P. 6

(www.subservientchicken.com;))

**Comparing Search Engines: Is Google or Yahoo More Cost Effective When Using
Sponsored Links for Small Companies?**

Kristin Kennedy, Ph.D.
Mathematics Department
Bryant University
Smithfield, RI
02917

Bonnie Brayton Kennedy, CPA
Controller
Kennedy, Incorporated
North Kingstown, RI

Introduction

Today it is imperative that all businesses, large and small, have an internet presence. With over 1 billion users in 2006 spending an estimated \$102.1 billion in online sales, it can hardly be ignored (Burns, 2007). For many small businesses, developing an effective web site that allows online transactions or drives customers to a physical location is often the easy part. The difficulty is getting seen by potential customers, and companies want to do everything they can to track this information. The key to being found by the right people at the right time lies with search engines. There are essentially two ways a customer will find a business site via a search engine, through an organic or a pay-per-click listing.

The organic listing is generated at no cost to the business owner. The search engine generates a ranked listing of sites relevant to the key word used by the customer. The sites listed first are those deemed most relevant by the search engine based on factors like site content, links, and current updates. Typically the organic listings are the bulk of the copy on the screen, and each search engine has their own algorithm for ranking the list of sites. Indeed this algorithm can often change.

Pay-per-click advertisements (PPC) are listed on the side, usually shaded or noted as "sponsored" links so that a customer will understand that the link was purchased by the web site business. This purchased listing is dichotomous to a listing that is organically ranked, and it puts forth a different message to the customer, not necessarily a negative message, but certainly different. The price paid per click is determined on an auction basis. For example, if a business wants the top spot for the term "gold ring", it will bid up to a certain amount "per click". Whenever the search engine displays an advertisement for a business web listing and a customer clicks on it, that action leads the customer to the business web site, and the company is charged an amount of money by the search engine company. Since this charge occurs each time a customer clicks through to a business web site and since this cost is changed regularly on an auction basis, it is obvious that this can be an expensive process, especially to a small company. Two questions come to mind. Is the PPC cost necessary to the business, or is the organic listing just as productive to the small business? Can it be shown that Google or Yahoo is a more cost effective search engine for a small company?

At first glance, it appears that the PPC option would be very productive for small businesses. Although it can be costly, sometimes it is the only way a company's web site will be seen by customers. Also it is difficult for small companies to compete with the giants in the organic click venue, because in the organic stream, the larger companies will generally land on top of the list. For example, a small manufacturer of gold jewelry will probably not get top billing over Tiffany's on the organic list, unless they spend a significant amount of time and money on developing their web site design and advertising. Furthermore, most web designers generally agree, that spending some money on PPC is well-spent money. It affords a company the opportunity to see where the company is placed on the sponsored link with the use of certain key words. Spending some money on PPC campaigns can help management design their web site to be optimized for organic search engine rankings. Once a company decides to commit to a

PPC program, then how does management decide which search engine is the most cost effective one for that particular company? This can be a crucial question to the small company working with a tight budget and a small staff.

The following paper examines a small manufacturing company's data from a PPC campaign with Google and separately with Yahoo. For seven campaigns, the company allowed the same amount of money in the budget with both search engines during the same time period. The management of the company had the "gut feeling" that they got more for their money with Yahoo, and the results highlighted in this paper tend to support that thesis, which was initially established with Kennedy and Kennedy (2007).

Statistical Review

Initially, the management of Kennedy Incorporated was interested in examining whether Google or Yahoo returned better average positions for keywords. The company had data from April to October of 2006 summarizing the 7 campaigns in Table 1. The campaigns were run on the exact same days with both search engines for comparison sake. July was omitted due to company vacation and closure. In reference to the last column of Table 1, Yahoo more often returned a better average position than Google; however there was not a huge range between the averages. An average position of 2 would be very good, and Yahoo returned an average closer to 2 than Google did. Only the two trials in June indicated that Google surpassed Yahoo. Moreover, the original raw data (not listed) seemed to indicate that the average position for each keyword that Kennedy Incorporated used had a significantly higher variance of position with Google. This is important information to the management. Only keywords that return a good position are of value to the company. They translate into visibility to the customer. If the average position of keywords had a greater variance with Google than Yahoo, then the company would be concerned that although the overall average positions were similar from month to month, within the month the company was losing visibility to customers by using Google.

Table 1: Summary of 7 Campaigns for Google/Yahoo

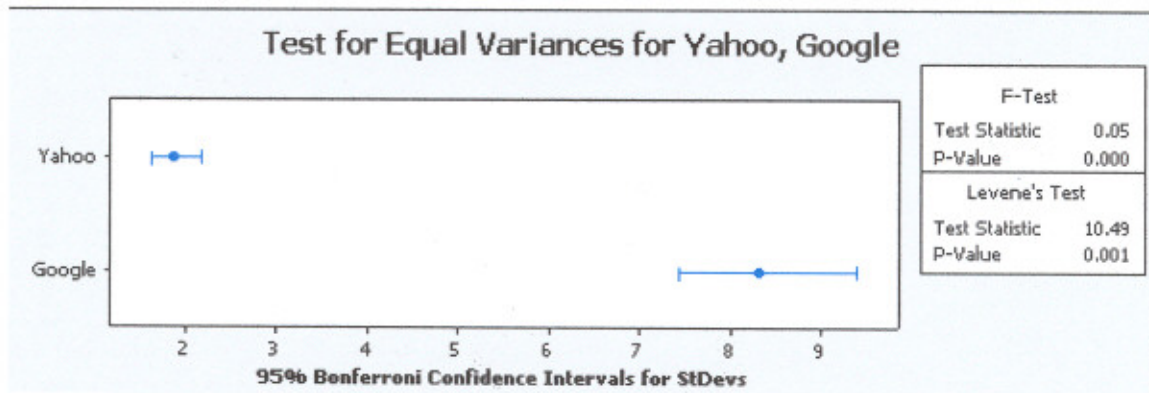
Campaign Period Google/Yahoo	Impressions Google/Yahoo	Click Google/Yahoo	CTR Google/Yahoo	Avg CPC Google/Yahoo	Cost Google/Yahoo	Avg Position Google/Yahoo
April Test	1,991/3,281	62/44	3.11%/1.34%	\$.79/\$.21	\$48.84/\$9.04	4.5/1
10Day June 4	104,370/9,624	275/200	.26%/2.08%	\$1.08/\$.28	\$298.35/\$56.61	4.5/1
10Day June 20	9,012/6,242	102/197	1.13%/3.16%	\$1.62/\$.68	\$164.99/\$134.79	2.7/3
June 21-30	8,668/5,909	102/222	1.18%/3.76%	\$1.47/\$.71	\$150.22/\$156.78	2/3
August 2006	17,082/4,694	109/180	0.64%/3.83%	\$1.40/\$.85	\$152.18/\$152.98	2.8/2
September 2006	10,776/4,071	62/178	0.58%/4.37%	\$1.37/ \$.85	\$85.08/\$151.62	3.1/2
October 2006	15,285/2,639	69/164	0.45%/6.21%	\$1.57/\$1.01	\$108.62/\$165.94	3.6/2

Each of the seven runs, highlighted in Table 1, was typically a small sample set. To try to ensure valid results, all of the data for Yahoo and Google was collected into one sample with two sets of average position data, one column for Yahoo and one for Google. The authors ran a quick graph of both the Google and Yahoo data. (Graphs not shown) Both the Google and Yahoo data were skewed to the right and uni-modal on the left

portion of the graph. The graphs did not suggest strict normality, but the data sets were large, and originally all the available data was used. The authors decided to use two parametric tests, but then also run a non-parametric test of medians to see if the results were much different. Three tests were performed on this data: 1) a test of variances, 2) a test of means, and 3) a test of medians.

All statistical tests were done in Minitab. We note p-values to determine significance and confidence intervals. To test if the variances were equal or not, we used the Levene test that does not assume normality. The p value was significant ($p = < 0.001$). Therefore the variances are not equal. Yahoo had a variance of 1.9 and Google a variance of 8.3, quite different, and the confidence interval for Google is much wider as noted in Figure 1.

Figure 1:



For the test of means, we used a 2 sample t test, assuming non-equal variances. Once again the p value was significant ($p = < 0.000$), and thus the two means are not equal. The mean for Yahoo average position was 3.3, while Google had an average position of 5.64. Noting that means are pulled to extremes with the data, we see that the Google average position is much higher, because there were some average positions with Google that were outliers. High average positions would be of little value to the company.

The last test performed was the non-parametric Mann Whitney test for medians. The non-parametric test was used to try to capture any discrepancy that could be found because of lack of normality. The two medians for Yahoo and Google were 3 and 3.5 respectively. The p-value was significant ($p = < 0.0034$). The two medians are not equal, and they are not pulled to extremes as the means are. However, we note that they are both skewed to the right, which is good. The values are skewed to the lower average position numbers.

The next two figures summarize the 95% confidence intervals for the mean and median for both Yahoo and Google.

Figure 2: Mean and Median 95% Confidence Intervals for Yahoo

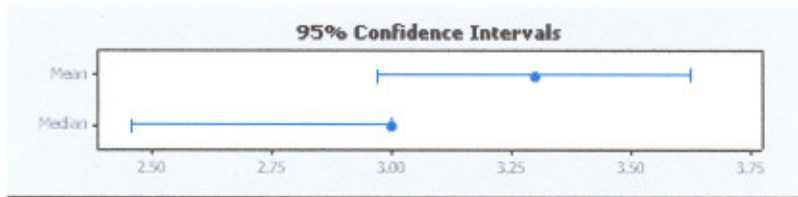
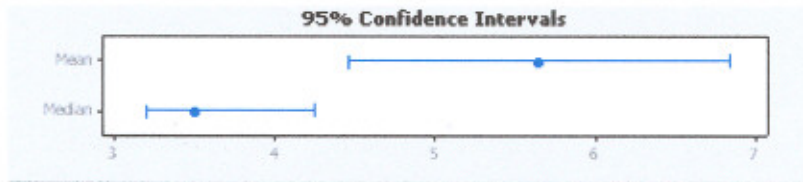
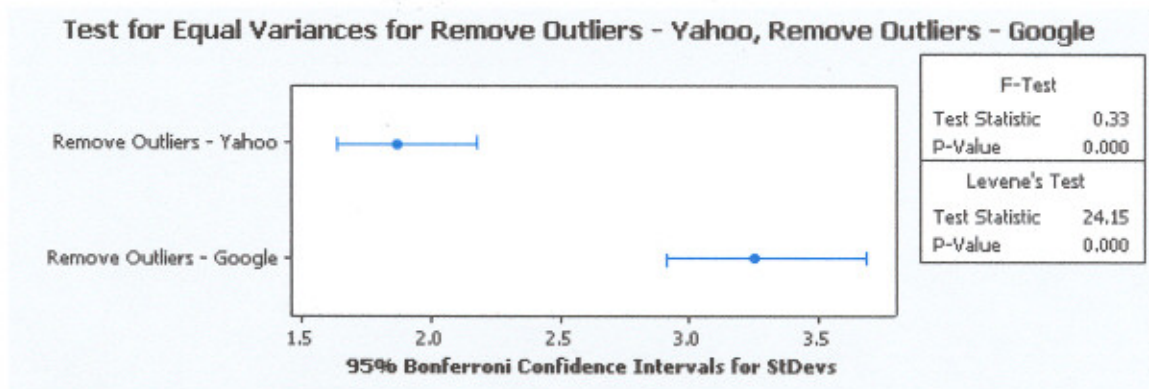


Figure 3: Mean and Median 95% Confidence Intervals for Google



In scanning the raw data and noticing that Google did have several outliers that would affect our results, we ran the same three tests again with the obvious outliers removed. This raw data is not listed in the paper. On the test for variances using the Levene test, a p value of 0.000 was returned; variances are still not equal. As a comparison to Figure 1, Figure 4 below shows some change in the confidence intervals, but not enough to change the significant outcome.

Figure 4: Test for Equal Variances with Outliers Removed



Using a t test for equal means, a p value of 0.000 was returned; the sample means were not equal. Finally, using the Mann-Whitney test for medians, the returned p value was 0.0099; the medians were still not equal. Thus even removing several obvious outliers, there was a big discrepancy with the average position concerning Google and Yahoo. For these campaigns, Yahoo is outperforming Google in returning significantly better average positions.

Continuing with the examination of some of the other data in Table 1, we see in column 2 that there were usually more impressions listed on Google. Impressions mean that the ad was displayed, however, an ad may have been so low on the list that the customer may not have even seen the ad (it may have been on the 8th page of results returned for a given search). In general the campaigns were set up to run for a specific time period, with the same dollar amount set for Google and Yahoo. For example in September, they set a budget of \$150 for each. However, Kennedy Incorporated reported that in September there were not enough clicks on Google to spend the allocated dollars for that period even though there were significantly more impressions on Google. Thus the high number of impressions did not help the sales of Kennedy Incorporated.

For each campaign in the test period, the CPC on Yahoo was significantly lower than Google. Furthermore, in all but one campaign (April test), the CTR was also significantly higher on Yahoo. In 5 out of the 7 runs, Yahoo had a CTR that was higher than 3%. Remember that the CTR is the number of clicks divided by the number of impressions. That suggests that Kennedy Incorporated received greater ROI from its advertising dollars spent on Yahoo, because more traffic was generated from the advertisements on Yahoo per dollar spent.

Conclusions

The management of Kennedy Incorporated has the impression that Yahoo is a better search engine for their company; they feel that they get a better ROI. This impression seems to be reinforced by the above statistics. Small companies must be very protective of every dollar spent. Kennedy, Incorporated wanted to step back from its internet advertising last summer to review their numbers. They now have two choices: 1) they can continue to work with both companies and collect more data or 2) they can spend their entire budget on Yahoo. As a continuance of this paper, we would like to examine some data from other small to medium sized companies. Does the data from other companies illustrate similar findings? Furthermore, in 2007 Yahoo restructured its search marketing bidding process. We would like to see whether the new format of Yahoo advertising will change the results for the company.

The real intent of the advertiser though is not simply to generate traffic, but to generate sales. So the question remains, does the higher CTR on Yahoo result in higher conversion to a sale. More data is needed.

Acknowledgements: The authors would like to thank Alan Olinsky for statistical consultation and guidance.

References

1. Burns, Enid. "E-Commerce Spending Up for 2007 Holiday and Year". January 5, 2007. http://www.clickz.com/showPage.html?page=clickz_print&id=3624408.
2. Kennedy, Kristin & Kennedy, Bonnie "An Examination of Search Engines for Small Companies: Is One Search Engine Better Than Another in a Highly Competitive Market?". The Proceedings of the 2007 Northeast Decision Science Institute, March 2007.

SURVEY DESIGN IN HEALTH CARE MARKETING RESEARCH: A BACK-TO-BASICS APPROACH

Amy Handlin
Dept. of Management and Marketing
Monmouth University
West Long Branch, NJ 07764
ahandlin@monmouth.edu
732-571-3659

ABSTRACT

Given the explosive growth of health care marketing and of the research that supports and critiques it, this paper argues that it is important to monitor – and to strive for continual improvement – in the survey design practices common in the field. In an exploratory examination of two criteria, the author finds that there is room for improvement both in sample design and in pretesting procedures.

INTRODUCTION

Since the turn of the decade, a sustained boom in health care marketing has caused it to surpass, in scope and influence, the marketing of many older product and service categories. For example, 2002 spending on direct-to-consumer prescription drug ads alone grew 5-10% -- despite a 9% drop in total advertising spending across all industries [1]. In the same year, the annual number of advertisements promoting both over the counter drugs and prescription medications was estimated at approximately 34,000 [2]. Fueled by an unprecedented combination of favorable demographic trends (in particular, the aging of the population), aggressive product/service development, and relaxed FDA regulations, this boom is projected to continue for the foreseeable future.

Indeed, while DTC prescription drug advertising is among the most controversial promotional practices, it is also among the most effective in terms of apparent influence with consumers. According to a 2003 survey by the FDA, 92% of patients had recently asked about an advertised prescription drug, “with 86% identifying the brand and 59% requesting a prescription for that drug” [3]. This degree of influence over their patients has long concerned physicians, who together with public health advocates have pressured the government to rein in pharmaceutical marketing. Thus far, the government has not acted – but the industry has voluntarily restricted its own promotions in a preemptive bid to head off more regulation [4].

Given these concerns – not to mention the potential for harmful consequences of making the wrong choice among many health care products – it is reasonable to expect health care marketers to uphold the highest standards for empirical research. In particular, since the majority of studies reported by industry marketers utilize some form of survey research, the design of these surveys should adhere to the best practices disseminated by survey practitioners overall. The purpose of this paper is to begin an examination of these practices.

BACKGROUND

There is precedent for training a critical eye on survey practices in consumer and marketing research. Jacoby and Handlin [5] noted that litigation in critical business matters, such as deceptive advertising and trademark protection, often relies upon surveys “to provide evidence of secondary meaning, genericness, likely confusion and the impact of disclaimers” [6]. In an extensive study of relevant court cases, they found that a not insignificant number of judges had largely dismissed survey evidence on the basis of inadequate survey design – specifically, because these surveys utilized nonprobability samples. As one judge wrote, “While nonprobability survey results may be admissible, they are weak evidence of behavior patterns in the test universe” [7]

Handlin and Jacoby pointed out that, notwithstanding courts’ criticisms, nonprobability samples are relied upon in the majority (indeed, over 90%) of empirical literature published in scholarly journals in the social and behavioral sciences. Clearly, there are well-designed, reliable nonprobability samples as well as poorly executed ones (the same can, of course, be said of probability samples). However, this debate in the courts has enduring value as a reminder that methodology matters.

METHOD

For the purpose of a preliminary review of survey design in published health care marketing research, the author focused on articles that appeared during a four-year period (2002-2005) in *Health Marketing Quarterly*. This journal was chosen because of the unusual degree to which it straddles both academic and practitioner research. The time period was chosen because it reflects the years of fastest growth in health care marketing (particularly for drug brands).

Of the 63 articles published during the time period, 51 were empirically-based studies. Two survey design criteria were examined in each of these 51 studies: (1) sample selection (probability or nonprobability?), and (2) whether questionnaires were pretested to identify problems prior to the survey. These criteria were selected from among the standards and best practices advocated by the American Association for Public Opinion Research and compiled by the Section on Survey Research Methods of the American Statistical Research Association. The choice of these particular criteria was made partly because of their broad applicability, and also because authors are more likely to discuss them than other criteria (such as whether alternative data collection methods were considered, or how a certain design balanced costs with potential errors).

The AAPOR/ASRA standards for sample selection and questionnaire pretesting are described, in part, as follows:

- (1) “In a bona fide survey, the sample is not selected haphazardly or only from persons who volunteer to participate. It is scientifically chosen so that each person in the population will have a measurable chance of selection...Virtually all surveys taken seriously by

social scientists, policy makers, and the informed media use some form of random or probability sampling...”

- (2) “High quality surveys and polls always provide adequate budget and time for pretesting questionnaires...Because it is rarely possible to foresee all the potential misunderstandings or biasing effects of different questions...it is vital for a well-designed survey operation to include provision for a pretest.”[8]

RESULTS AND DISCUSSION

As shown in the table, about 37% of the 51 empirical studies published by HMQ between 2002-2005 were based on probability samples. Well over half (62%) of the studies were based on nonprobability samples. One of the studies reviewed lacked sufficient detail to categorize the sample.

Survey Design Criteria Results

Issue	Survey-Based Articles (#)	Sample Used			Pretest	
		Probability	Nonprobability	Indeterminate	Yes	Not reported
2002(1)	3	2	1		1	2
2002(2)	4	1	3		2	2
2002(3)	2	1	1		2	--
2002(4)	4	--	3	1	1	3
2003(1/2)	5	3	2		2	3
2003(3)	4	2	2		2	2
2003(4)	4	1	3		--	4
2004(1)	2	2	--		2	--
2004(2)	4	1	3		2	2
2004(3)	4	2	3*		2	3
2004(4)	2	--	2		1	1

2005(1)	3	2	1		--	3
2005(2)	3	--	3		--	3
2005(3)	4	1	3		1	3
2005(4)	3	1	2		1	2
Totals	51	19	32	1	19	33

*One study utilized two samples, one probability and one nonprobability.

Likewise, some type of pretesting procedure was reported in about 37% of the 51 studies. In the remaining studies, there was no documentation of whether or not pretesting was done.

With respect to sample design, these results are encouraging. Clearly, significant efforts are being made by health care marketing researchers to utilize probability samples when possible and practical. At the same time, the strong predominance of nonprobability samples means that there may be room for improvement. Also, it serves as a reminder to researchers that they are obligated to fully explain their reasons and justification for choosing one type of sample over another.

The findings regarding pretests are more troubling. While 37% of the studies document some use of a pretest, the remainder fail to mention this criterion at all. This may indicate the presence of two related but separate problems. First, it is possible that the majority of researchers are not conducting pretests. Second, even those who are pretesting may not appreciate the importance of documenting how or why it was done.

LIMITATIONS

Because this examination is based on empirical articles appearing in only one journal, it may have limited generalizability. Indeed, the author recognizes that these findings may reflect editorial preferences as much as trends among researchers in the health care marketing field. However, this preliminary exploration can serve, at a minimum, to draw attention to low profile but potentially serious methodological issues. Among both scholars and practitioners, these issues merit broader and deeper analysis.

RECOMMENDATIONS

As markets continue to expand in the health care industry and competition heats up, especially among prescription drug brands, it is likely that research will increase in importance. This means that researchers will have a steadily growing responsibility – along with more opportunities -- to do their best work.

This preliminary exploration should be expanded by examining survey design in studies published by other influential journals in the specialized field of health care marketing. It would also be helpful to compare the design of studies in these journals with the design of health care studies published in journals of general marketing research.

This paper, if its findings are supported by further work, should be of interest – indeed, of concern – to both scholars and practitioners. Its “back-to-basics” message is a timely reminder of those fundamental principles that determine the quality of all empirical research.

REFERENCES

- [1] Dickinson, James. “DTC Report: FDA Survey Data Supports DTC,” *MM&M*, 2002 (June), 32 and 34.
- [2] Adams, Chris. “FDA Inundated Trying to Assess Drug Ad Pitches,” *Wall Street Journal*, 2002 (March 14), B1.
- [3] *Broadcasting and Cable*, “FDA Study Supports DTC Ads,” 2003 (Jan.27), 19.
- [4] Schiffman, Leon and L. Kanuk. *Consumer Behavior*. Upper Saddle River, NJ: Prentice-Hall, 2007.
- [5] Jacoby, Jacob and A. Handlin. “Non-Probability Sampling Designs for Litigation Surveys,” *The Trademark Reporter* 1991 81(2), 169-179.
- [6] Jacoby and Handlin, *Ibid* 169.
- [7] American Home Products Corp. v. Barr Laboratories Inc.. 656 F. Supp. 1058, 3 USPQ2d 1194 (D NJ 1987).
- [8] American Association for Public Opinion Research. “Standards and Best Practices for Survey and Public Opinion Research,” www.aapor.org 2007, 2-4.

DESIGNING FOR DELIGHT AND DISASTER: ENHANCING INTRINSIC AND SITUATIONAL INVOLVEMENT THROUGH WEB DESIGN

Dr. Laura Lally

Associate Professor

Hofstra University

Business Computer Information Systems Department

Hempstead, NY 11549-134

516 463 5351

acslhl@hofstra.edu

Keywords: Consumer involvement theory, customer value model, E-Commerce, high engagement, high intrinsic involvement, high situation involvement, low engagement, Web site design.

ABSTRACT

This study will draw upon the *theory of consumer involvement* and explore its implications in designing appropriate Web based applications. The theory will be extended to create a four part model—High Intrinsic Involvement, High Situational Involvement, Low Involvement, and High Engagement, based on different ways of providing of customer value. Research propositions for empirical testing of the model and implications for Web design will result from the analysis.

INTRODUCTION: E-COMMERCE ENTERS A NEW DECADE

After an exciting and tumultuous first decade, E-Commerce has experienced a wide range of successes and failures. The use of the World Wide Web has penetrated almost all forms of business and most levels of society, and has become a regular part of the lives of business managers and individuals. Because of the wide acceptance of the Web, more powerful, robust infrastructures are emerging, and standards are being developed to support more comprehensive and complex applications. Assured of a better functioning infrastructure and the potential of a large customer base, designers of next generation Web based businesses still need to know which applications will provide value to customers and profits to the businesses that create them.

The Internet has enabled an increase in *information richness*—the complexity and content of Web based information (Evans and Wurster, 1999) now made possible by multimedia. The Internet has changed the traditional tradeoff between *information richness* (the complexity of the message) and *information range* (the number of people who receive the message), making richer information available on a global basis. The Internet has also enabled an increase in *information density*—the total amount of information available to all participants in a decision as well as the potential for improving its currency and accuracy, creating the potential for better quality decisions

(Laudon and Traver, 2007). This paper addresses the issue of how Web sites can use these emerging information qualities to better satisfy customer needs.

This paper proposes a *customer value model*, which focuses on the *value to the customer* provided by an E-Commerce application, and argues that it is only by providing value to customers that an E-Commerce site can generate revenue to the creators of the Web based businesses. This study will draw on the theory of *consumer involvement* from the field of Marketing and develop a model of customer value based upon the customers' underlying involvement with the product or service offered by the Web site. This approach is in contrast to earlier studies that focused on the customers' involvement with Web use in and of itself. The result of the analysis will be a four-part model of customer value that will have implications for Web site development and design, the appropriate business model for making the site profitable, and the need for further basic research for providing value to customers.

THEORETICAL PERSPECTIVE: INVOLVEMENT THEORY

This study will draw on *involvement theory* from consumer behavior research to inform the analysis. A consumer's involvement with a product can be defined in terms of, "the degree of personal relevance that the product holds for the consumer. Under this definition, high involvement purchases are those that are very important to the consumer and thus provoke extensive information processing," (Schiffman & Kanuk, 1999, p. 186). Andrews, Durvasula, and Akhter (1990) define involvement as "an individual, internal state of arousal with intensity, direction, and persistence properties." Consumer behavior further distinguishes between *intrinsic involvement*, which occurs when an individual finds the product personally relevant, inherently meaningful, and enjoyable, and *situational involvement*, which occurs when an individual is in a risky situation such as making an expensive purchase, or facing a life threatening emergency Celsi and Olson (1988). *Low involvement* products tend to be commodities which consumers find neither fascinating nor risky. Although a number of studies (Venkatesh, 1999, Agarwal & Karahanna, 2000, Novak & Hoffman, 2000) addressed user's involvement with computer use and Internet surfing per se, Lally (2003, 2004) first addressed the idea of studying users' involvement with the underlying product or service as a determinant of Web site success. Lally, argued that the typical Internet user does not surf the Internet because they find computers or the Internet inherently fascinating but because the Internet: 1) allows them to interact with something they do find fascinating, 2) helps them gather information for dealing with risky situations, or 3) allows them to conduct routine transactions in a convenient and efficient manner. Her results indicate that users seeking serious information versus entertainment require fewer of the immersive features provided by multimedia, but require more timely, accurate content.

HIGH INTRINSIC INVOLVEMENT: PROVIDE HEDONIC VALUE

Areas in which consumers develop high degrees of intrinsic involvement include music, art, films and sports, and other forms of entertainment where consumers develop intense personal interest. These individuals become knowledgeable about a field and

enjoy discussing it with others. High involvement customers find visiting retail outlets that let them experience their area of involvement inherently enjoyable, even if they are not planning an immediate purchase. This intrinsic pleasure in high involvement shopping is called *hedonic value*. Customers willingly devote time and attention to learning about the details of the product and develop sophisticated levels of knowledge. They often lose track of time when immersed in their area of involvement (Agarwal and Karahanna, 2000). They may also experience “cognitive absorption” a phenomena that occurs when the whole of an individuals consciousness is absorbed in an experience their area of involvement, making interruptions and distractions very unwelcome (Agarwal and Karahanna, 2000; Carver and Turoff, 2007; McCoy Everard, Polack, Galletta, 2007). When high involvement customers do make purchases, they are willing to pay a premium for products, services and experiences whose characteristics suit their personalities and satisfy their tastes. Their devotion to their area of involvement continues to endure over time as their knowledge and experience deepens. They enjoy enriching their knowledge and/or enhancing their performance and the range of their experiences in the area. High intrinsic involvement goods are typically sold in retail outlets where they are displayed in an attractive and involving manner to increase hedonic value. Sales representatives need to be well informed about product features. Educating potential customers and interacting with already knowledgeable customers is part of their job. Because these products tend to have a strong visceral component, customers are often given the opportunity to interact physically with the product before making a purchase.

In E-Commerce, Web sites targeting consumers with high intrinsic involvement must confront the limitations of the on-line shopping experience. Consumers cannot physically experience the products as completely as when they are in a retail outlet. As bandwidth limitations become less of a constraining factor, however, Web designers can use multimedia to compensate for the lack of physical presence. Lui, Arnett, and Litechy (2000) suggest that Web designers should aim to create hedonic value for users by making sites playful and engaging. High quality graphics, sound, and animations can provide virtual immersive experiences. On-line experts can provide advice about products. User groups can provide the enjoyable experience of discussing the products with others. High end auction houses have found, for example, that customers will purchase very expensive works of art over the Web if they can view details of the artwork through high quality graphics, and interact with art experts to become more informed about their potential purchase.

Because high intrinsic involvement addresses an enduring personal interest on the part of customers, a *subscription* revenue model is most likely to be successful. Users are likely to return to sites over and over again, just as they look forward to new issues of magazines on their favorite subjects. Therefore, content must be continually refreshed to provide new experiences with the desired product.

The following research propositions emerge:

Proposition #1: Individuals with high product involvement will prefer web sites with high information richness.

Proposition #2: The subscription model will be most appropriate for high product involvement.

HIGH SITUATIONAL INVOLVEMENT: IMPORTANT BUT NOT INHERENTLY ENJOYABLE

Unlike consumers with high intrinsic involvement, consumers with high situation involvement do not find their subject inherently enjoyable. Instead they invest time and attention because they are in a situation that is important, and risky. High situation involvement occurs when individuals are buying a home, seeking medical information when they or their loved ones are ill, or facing a natural or man made disaster. Individuals in these situations face a complex, unstructured task which they need to complete successfully, and so give it their full attention. The successful completion of the task can lead to positive results such as good physical or financial health, or a safe and secure environment that will lead to personal happiness, but individuals do not find the situation itself inherently enjoyable.

When facing high situation involvement scenarios, individuals want access to essential facts to help them make important risky decisions. They willingly engage in extensive information processing to help them make the right decision. Providing assurance and minimizing risk and anxiety is critical in high situation involvement. When a customer is making a risky decision, reducing the information asymmetry between the buyer and seller provides value. For example, a mashup of mapping and crime data can show potential homebuyers if the house they are buying is in an area with many home break-ins—a fact not likely to be revealed by the seller or real estate agent unless the number of crimes is very low. For managers of Web sites for high situational involvement scenarios, involvement theory suggests that the emphasis be put on informing, rather than entertaining the consumer. Immersive, exciting multimedia features, or information not relevant to the task at hand are wasted and even unwelcome unless they address the consumer's need for knowledge. Interruptions again are seen as intrusive and unwelcome (McCoy Everard, Polack, Galletta, 2007).

For consumers making one-time important decisions, a fee for content business model would most likely be the most profitable. Unlike the case with high intrinsic involvement, high situation involvement does not represent an enduring interest in a product, so users would be less likely to want to subscribe on a long term basis. These individuals do not want to continue increasing their knowledge and experience in the area, once the task is complete. However, in the case when there is an important decision that is made repeatedly, a subscription model would be most appropriate.

The following research propositions emerge:

Proposition #3: Individuals with high situational involvement will prefer Web sites with greater information density.

Proposition #4: The fee for content model is most appropriate for Web sites for high situation involvement if the decision is a one time decision. A subscription model is most appropriate if it is a repeated decision.

LOW INVOLVEMENT: MAKE IT QUICK AND SIMPLE

Low involvement products, in contrast, do not personally engage the user. They are neither intrinsically valuable nor do they involve complex and risky decisions. Cash from automatic teller machines, soda from vending machines, and office supplies are examples of products that customers seek out of necessity, not because they find the shopping experience inherently enjoyable. Low involvement product shopping, therefore, provides less hedonic value than shopping for high involvement products. Unlike high purchase involvement products, purchasing these products is neither expensive nor risky. Customers, therefore, do not need to a great deal of information to help them choose a product. Customers prefer “limited information processing when the purchase is of low personal relevance” (Schiffman & Kanuk, 1999, p. 186). Involvement theory suggests that an efficient purchasing process is the quality most valued for these products. Kaufmann and Lally (1994) found, for example, that for automatic teller machines, convenient location and easy access dominated design features in attracting customers. For low involvement purchases, customers should not be distracted by information irrelevant to their purchase or overloaded with unnecessary details.

A fee for transaction business model would be most appropriate. The goal of these sites is to decrease user’s transaction costs and users with heavy demands on their time and attention are frequently willing to pay a premium to minimize the time and effort required of them to conduct routine tasks. Usability testing upfront to make the site easy to use is of critical importance, but once the site is well designed frequent changes and updates are likely to confuse, rather than provide value to the customer. Examples of these sites include grocery ordering and electronic banking. Although ATM use and online banking for routine bill paying has been widely adopted, the number of bank branches continues to increase. This analysis suggests that when customers are making routine transactions, they are willing to use automated tools, but want to interact with human beings when making high situation involvement financial decisions.

The following research propositions emerge:

Proposition #5: Individuals with low product involvement prefer Web sites with less information richness and information density.

Proposition #6: The transaction fee model is most appropriate for low product involvement Web sites.

PASSIONS VERSUS PASTIMES: EXTENDING THE MODEL

An examination of E-Commerce applications such as downloads of comedy shows to cellphones, videos posted on uTube, MySpace, and computer gaming indicates that there exists a fourth category of applications. *These applications provide high intrinsic interest to the consumer, but are not of enduring personal importance.* “Cool” consumer based applications aimed at killing time in a pleasant and engaging matter have a significant number of customers and are likely to continue to do so in the next decade

as portable devices are developed that support these applications. This study proposes to extend the theory of consumer involvement to include applications that provide this form of value and will use the term High Engagement to address these types of applications. High Engagement: High Enjoyment, Low Importance.

These sites should provide a media rich, engaging interface and charge for the service they provide consumers with a pleasant means of passing the time, either through connect time or usage. As newer “cooler”, more exciting forms of high engagement applications become available, customers are likely to switch to them, hence the need for constant innovation. Because the degree of personal relevancy of a particular product is low, customer loyalty is likely to be fleeting. Since the goal of these sites is to provide entertainment, the analysis suggests that high engagement sites may successfully use a fee for use or advertising business model.

The following research propositions emerge for high engagement sites:

Proposition #7: Individuals with high engagement will prefer high information richness and low information density.

Proposition #8: The connect time or advertising revenue models are most appropriate for high engagement Web sites.

Table 1 summarizes the categories suggested by the theory of consumer involvement.

Table 1.

	Value to Customer	Revenue Model
High Product Involvement	Immersive interaction—high information richness and density.	Subscription
High Situation Involvement	Timely, accurate information—high information density.	One time fee for one time decision. Subscriptions for repeated decision.
High Engagement	Media rich interaction—high information richness and low information density.	Fee for connect time or advertisements.
Low Involvement	Quick efficient transactions costs—low information richness and low information density.	Fee for transaction.

References upon request.

CHARACTERISTICS AFFECTING THE ABANDONMENT OF E-COMMERCE SHOPPING CARTS – A PILOT STUDY

Jason Coppola, Bryant University, (203) 496-3234, Jason.Coppola@quinnipiac.edu
Kenneth J. Sousa, Bryant University, (401) 232-6334, ksousa@bryant.edu

ABSTRACT

The number of sales transactions and volume gained through web-enabled systems has steadily increased over the last few years. However, consumers continue to abandon the electronic shopping carts associated with these systems. This research study seeks to explain why electronic consumers abandon electronic shopping carts using categories as a means to distinguish between various reasons associated with behavior. This study surveyed 153 college students using questions associated with website design, security, social/psychological and website features to determine the effect on electronic shopping cart abandonment. The results confirm that there are statistically significant factors that lead to electronic shopping cart abandonment. The study also provides suggestions for website designers to decrease abandonment as well as proposing areas for future research.

INTRODUCTION

Since 2000, e-commerce sales have been growing at a steady rate (Meeker, Pitz, and Fitzgerald, 2004; shop.org, 2004). As e-commerce sales continue to expand and more people use the Internet as a shopping channel, issues have arisen with the usage and abandonment of electronic shopping carts. Electronic shopping carts can be defined as electronic commerce user-interface for the customer to shop at online stores.

Shopping cart abandonment is defined as, a user leaving a website after they have placed items into their cart without purchasing those items. The most successful e-commerce websites convert only 8 percent of their online visitors to paying customers. Most firms are able to convert only 2 to 3 percent (Goldwyn, 2003). Research has shown that two-thirds of online shoppers have filled their electronic shopping cart but exited at the check-out stage without completing a purchase (Gurley, 2000; Rewick, 2000). Even though e-commerce is a thriving industry, these conversion numbers are very poor.

This research study attempts to gain insight on why consumers, specifically college students, are willing to place items into their electronic shopping cart, but fail to complete the order.

LITERATURE REVIEW

Understanding the factors that influence consumers to abandon electronic shopping carts has been the goal of e-commerce research since the late 1990s. Even though there have been studies since the late 1990s, there are very few that focus on shopping cart abandonment. Cho (2006) determined that cart abandonment is more closely related to overall hesitation; not to hesitation associated to payment. The study also found several reasons that college students are overall more likely to hesitate when purchasing a product or service online including race, attitude toward online shopping, quality-conscious, buyer's regret and time pressure.

Value-consciousness is the strongest predictor of overall shopping hesitation. The study asserted that college students are more likely to remove product items off their online shopping carts based on several factors including time on the Internet, confusion, negative past experience, price comparison, reliability of website and depth of personal information required to complete online purchase.

In addition, comparison shopping is the strongest predictor of shopping cart abandonment (Cho, 2006). Other studies have focused on factors that affect attitudes towards retail websites (Elliot, 2005). This study found that as expectations of online shoppers rise, their satisfaction with online retailers decline. The results of their study assert that five website factors (ease of use, product information, entertainment, trust, and currency) affect consumer attitude towards a retail website. The study also suggested that retail websites should emphasize site factors that best suit the involvement/experience profile of their primary users (Elliot, 2006).

Another study focused on creating value for online shoppers (Lee, 2004). This study identified two types of online shopping values: utilitarian value (including price savings, service excellence, time savings, and selection dimensions) and experiential value (including entertainment, visual, escape, and interaction dimensions). The results of the study indicated that Internet shopping does invoke various types of shopping values. It also asserted that both utilitarian and experiential value positively affected customer satisfaction (Lee, 2004).

McCloskey (2006) focused on the importance of ease of use, usefulness, and trust to online consumers. This study examined electronic commerce participation by older Americans. This study used and modified the Technology Acceptance Model (TAM) to examine the impact attitudes concerning ease of use, usefulness and trust had on electronic commerce usage. Usefulness and trust were found to have a positive, direct effect on usage. Ease of use had significant impacts on usefulness, while trust had a significant impact on both ease of use and usefulness (McCloskey, 2006).

Ha (2004) extended Roger's innovation decision process to study how general innovativeness is related to Internet apparel shopping. Using college students as the survey population, the results asserted that general innovativeness is related to Internet usage for information search, but not actual purchases. The study also found that gender is related to general innovativeness, because female students were more likely to use the Internet for information search for apparel products.

A research study by Li (2005) focused on shopping cart abandonment at retail websites and developed a multi-stage model of online shopping behavior. This study applied their model to click-stream data from www.barnesandnoble.com. The study identified three set of factors that affect shopping cart abandonment: (1) consumer's unobserved purchase intent that influences navigation orientation, (2) exposure to hyperlinks, marketing mix and promotional stimuli on web pages while browsing a website that can change navigation orientation at any point during the session, and (3) comparison shopping activity that affects consumer's progress through the shopping process.

Lastly, Wolfinbarger (2001) reviewed what types of shoppers are interested in e-tailing. The study found that goal-oriented or utilitarian shoppers, who are described as task-oriented, efficient, rational, and deliberate, desire to purchase what they want quickly and without distraction. This research suggests that goal-oriented shoppers are interested in e-tailing because of four specific attributes: convenience and accessibility; selection; availability of information; and lack of sociality.

PROPOSED RESEARCH MODEL

This research study proposes a model that includes four factors which affect electronic shopping cart abandonment. Those four factors are: 1) website design, 2) security, 3) social/psychological factors, and 4) website features. The measures for this study were compiled from the proposed research model described above. The literature review provided the basis to identify construct definitions and any prior measures. Seven-point Likert scales were assigned to each construct as outlined.

Constructs

Website design is how complicated the website is designed, the time required to download necessary product information, and if the site had a fun and entertaining feel. Three questions were designed to understand how electronic shoppers feel towards website design using a strongly disagree to strongly agree scale (Cho, 2006).

Security is defined as features that the website has implemented to protect its customers personal information. Four questions developed the security portion of the survey (strongly agreed to strongly disagree scale). These questions helped to understand the extent to which e-shoppers care about the security features of a site before they purchase from that site (Cho, 2006; McCloskey, 2006; Elliott, 2005).

Website features are special characteristics that the website has which increases its value against competitors; including shipping charges and return policies. Six questions were used to compile which features of a website entice e-shoppers to remain on a website and complete a purchase using a strongly disagree to strongly agree scale (Cho, 2006; Elliott, 2005; Lee, 2004).

Social/psychological factors are characteristics that the consumer personally considers before they purchase a product. These factors are not affected by the website design. Five questions from the survey were developed toward understanding how an e-shopper's social and psychological factors affected whether a purchase was completed using a strongly disagree to strongly agree scale (Cho, 2006; Wolfinbarger, 2001).

Abandonment is defined as a user leaving a website after they have placed items into their cart without purchasing those items. Three questions were used to understand how many times the respondents have abandoned electronic shopping carts in the past. These questions also helped to determine whether they abandon their cart to purchase at a later time or completely as well as whether they abandoned all of the items in the cart or only a few of the items (partial order) using a never to always scale (Cho, 2006).

The following hypotheses will be utilized to determine the relationship between these factors and shopping cart abandonment. The hypotheses are as follows:

Hypothesis	Description
1	Website design does not affect shopping cart abandonment.
2	Website security does not affect shopping cart abandonment.
3	Social/Psychological factors do not affect shopping cart abandonment.
4	Website features will have no effect on consumers abandoning shopping carts.

Survey Implementation

College students were selected as the survey population due to their use of technology, significant use of the Internet and are more likely to abandon products online (Cho, 2006; Ha, 2004). A web-based survey was developed consisting of twenty-six questions categorized into six categories: population, abandonment, website design, security, website features and social/psychological factors.

RESULTS

The survey was distributed to 1,000 students. A total of 153 responses were received resulting in a 15.3% response rate. There were 60 male responses, 54 female response, and 39 responses that did not identify gender. The results were analyzed using factor analysis which resulted in four significant factors. After completing and analyzing the results of the factor analysis, a linear regression analysis was completed to determine which of the four factors loaded onto the dependant variable (abandonment). The model's four factors explain 60.09% of the total variance. The significant factors were selected based on variables with communalities more than .50 significance (Hair, 1998).

The factor analysis was completed using the four factors outlined in the literature review. Each of the survey questions resulted in communalities which are greater than .50 under the rotated factor matrix. The following table summarizes the results of the factor analysis:

Factor	Construct	Question (factor value)
IV-1	Website design and security	<ul style="list-style-type: none">• Complicated checkout process (.607)• Downloading product information too long (.587)• Checkout required too much personal information (.787)• Site was unreliable (.785)• Lack of confidence in secure credit card transactions processing (.724)• Concern for misuse of personal information (.640)
IV-2	Website features	<ul style="list-style-type: none">• Site is well-organized and easy to navigate (.701)• Useful search engine to find information (.718)• Site contained useful graphics of products (.874)• Products clearly described (.731)
IV-3	Social and psychological	<ul style="list-style-type: none">• Needed additional advice for purchasing decision (.688)• Avoiding buyer's remorse (.577)• Possibility of a new product introduced after purchase (.673)
IV-4	Website functions	<ul style="list-style-type: none">• Website has a difficult return policy (.627)

A linear regression analysis was completed to determine which of the four factors loaded heavily onto the combined dependent variable which included four questions. The dependent variable construct was created by averaging the values of the four questions. The dependent variable was then tested against the four factors (IVs). The four factors were created by multiplying each question by their communality score in the rotated factor matrix and then added that result to the next significant question in that factor. The dependent variable constructs were compiled as follows:

Factor	Construct	Question (factor value)
DV	Abandonment	Abandoned the shopping cart ... <ul style="list-style-type: none"> • ... delaying the purchasing decision • ... refusing the entire cart purchase • ... items for that time and in the future.

This statistical test identified two significant independent variables which load heavily onto the combined dependent variable. The two independent variables were factor 1 and factor 3. The regression model accounted for an R-squared of .127 concluding that 12.7% of the total variance is explained by the independent variables. The significance of factor 1 was .000, which means that it is very highly significant. The significance of factor 1 and factor 3 combine was .011; also highly significant.

Factor 3 calculated a standardized beta of .294 while factor 1 and factor 3 combined resulted in a standardized beta of .210 and .239 respectively. The standardized regression coefficient, also known as beta, allows for a direct comparison between coefficients as to their relative explanatory power of the dependent variable. The estimated regression coefficients are used to calculate the predicted values for each observation and to express the expected change in the dependent variable for each unit change in the independent variable(s).

After the initial regression using the factors calculated from the combined questions was completed, the individual questions specific to the significant factors were analyzed through a linear regression with the combined dependent variable. The regression identified the questions relating to the website's unreliability and buyer's remorse as significant. The regression model accounted for an R-squared of .145. The significance of the unreliability construct was .000 while the combined unreliability and buyer's remorse was .002; both highly significant. Unreliability calculated a standardized beta of .289; unreliability and buyer's remorse combined resulted with a standardized beta of .302 and .248 respectively.

FINDINGS AND CONCLUSIONS

Limited research on the rationale relating to shopping cart abandonment has been limited. This research study considered the impact of four constructs (security, website design, website features and social/psychological) on online shopping cart abandonment. The results found that security, website design, website features, and social/psychological factors each affect shopping cart abandonment.

When linear regression was used to determine the variance explained by the factors only 12.7% was accounted for. The analysis clearly indicates that two factors, website design/security and social/psychological characteristics, are significant predictors. Website design appears to have a significant effect on the abandonment of online shopping cart. If a website has a simple checkout they are more likely, according to this study, to retain a customer's purchase. Also, if the time to download product image was lengthy, consumers were likely to abandon their shopping cart. Website designers should focus more on these two features when designing a website to increase online shopping purchases.

Security features on websites also seem to have a significant effect on the abandonment of online shopping carts. As security features increase, website abandonment decreases. Therefore, increased emphasis on security features should increase customer purchasing behavior.

Website designers should also consider consumers' social/psychological characteristics. Based on this study, these factors load heavily onto online shopping cart abandonment. Consumers may want to avoid buyer's remorse. Consumers are also concerned that improved products may enter the market soon after purchase and therefore more willing to abandon shopping carts. Lastly, consumers need assistance when completing purchasing decisions. If they are unsure about purchasing a product and can seek assistance while online, they are less likely to abandon their shopping cart. Understanding a potential customer's social/psychological factors can assist designers implementing websites which alleviate customers' concerns.

This study was also able to identify that a website's reliability and a consumers wanting to avoid regrets has significance with shopping cart abandonment. Meaning as website unreliability increases shopping cart abandonment also increases. Also if a consumer is concerned about avoiding regrets they are more likely to abandon a shopping cart.

This study was able to identify three areas that influence online shopping cart abandonment; website design, security and social/psychological characteristics. The results of the research indicate that college students are concerned with the simplicity of a website's checkout, the speed at which the website completes its download, the security of the website and social/psychological concerns.

Even though through factor analysis the study was able to find four factors loading onto online shopping cart abandonment, the results could not be strongly supported. The low R-squared does not provide a strong influence, but may provide a foundation to future research by probably expanding the population sample and increasing the response rate. Collaterally, some significance in other website design and website features were found to affect shopping cart abandonment but this study was unable to prove how they affect abandonment.

LIMITATIONS AND FUTURE RESEARCH

There are a few limitations to this study. First, the sample size was sufficient to complete the statistical tests but the data could be more predictive if the sample size was larger. The research could only identify three factors that influence shopping cart abandonment. Also, the regression R-squared value resulted in a low value (12.7%) of the total variance explained.

Future research could be focused on how website design and features can be enhanced to limit the shopping cart abandonment rate. Also, future research should focus on how a website security features can be explicitly displayed and enhanced to ease customers fears about how their personal information will be used. An additional area of future research could focus on how specific social and psychological factors explain shopping cart abandonment. Lastly, future research could attempt to explore the effect of an organization's online shipping and return policies on shopping cart abandonment and any effects on whether these policies were displayed more predominately on their websites.

REFERENCES

References and tables furnished upon request of author.

SPAM E-MAIL MARKETING OF SEXUAL PERFORMANCE PRODUCTS

Joshua Fogel, PhD, Department of Economics, Brooklyn College of the City University of New York, 2900 Bedford Avenue, 218A, Brooklyn, NY 11210, (718) 951-3857, joshua.fogel@gmail.com

Sam Shlivko, Department of Economics, Brooklyn College of the City University of New York, 2900 Bedford Avenue, 218A, Brooklyn, NY 11210, (718) 796-2344, samshlivko@gmail.com

ABSTRACT

We study spam e-mail use among those with and without sexual performance problems (n=200). Questions inquired about receiving, opening, and purchasing from spam e-mail regarding sexual performance products. Those who had sexual performance problems were more likely than those who did not have sexual performance problems to open and purchase from spam e-mail regarding sexual performance products. Spam marketing e-mails for sexual performance products have a receptive audience. We are not advocating sending spam e-mail. However, clearly there is a demand for e-mail marketing of sexual performance products and this should be done through ethical and appropriate marketing channels.

Keywords: direct and mail marketing, direct-to-consumer marketing, e-commerce, electronic mail, Internet

INTRODUCTION

Internet spam e-mail constantly bombards our inboxes. A survey conducted by the Pew Internet and American Life Project in 2007 reported that most e-mail users are bothered by spam with 18% believing that it is a big problem and 51% believing that it is annoying [1]. In another study, 81% of individuals “strongly disliked” spam, with almost all the rest of the sample “disliking” spam [4]. Health-related messages consist of 32% of spam e-mail [3].

Spam e-mail takes marketing to an extreme, with many businesses and consumers considering it an unsavory marketing approach. However, clearly there are consumers who respond to these spam e-mail solicitations and that is one of the reasons why certain businesses send spam e-mail. Age is related to consumer purchases from spam e-mail. Older retirement-age individuals are more likely to purchase from solicitations in spam e-mail than college-age young adults [4]. There is no precise understanding of the extent of this target market who purchase from spam e-mail. Grimes [5] quotes the Direct Marketing Association who report that 36% of e-mail users purchase a product from spam e-mail [6], while others report a much lower percentage where only 8% of e-mail users purchase from spam e-mail [9].

Spam e-mail range in the variety of products and services offered and a portion of this spam e-mail appeals to the health concerns of individuals. This is a potentially large group of individuals, as the Pew Internet and American Life Project consistently report from data

collected in 2002, 2004, and 2006 that 80% of United States Internet users look on the Internet for health information [2]. One common type of spam e-mail involves remedies for sexual performance problems. This can include offering for sale natural remedies or the prescription medications of Viagra or Cialis [3] which are typically used to treat erectile dysfunction, which is a leading cause of sexual dysfunction in men.

In the United States in 1997 the Food and Drug Administration (FDA) allowed pharmaceutical companies to include the brand name of their products in their advertisements. This resulted in an explosion of direct-to-consumer advertising of prescription drugs from 12 million dollars in 1998 to 2 billion dollars in 2000 [10]. Spam e-mail is an extension of this direct-to-consumer advertising, although much of the origin of the advertising to United States consumers does not necessarily originate in the United States [3].

The objective of this study is to determine if the presence of sexual performance problems is associated with receiving spam e-mail containing sexual performance product solicitations, opening of spam e-mail containing sexual performance product solicitations, and purchasing from spam e-mail containing sexual performance product solicitations.

METHOD

Participants and Procedures

Participants were 200 students from a 4-year undergraduate commuter inner city college. Of 212 students approached, 200 completed surveys for a response rate of 94.3%. The sample was a convenience sample. Participants were approached in classrooms, the school cafeteria, the library, and other public places. They were asked to complete an anonymous survey on spam e-mail. Data collection occurred during May 2007. The survey was exempt from Institutional Board Review and was conducted in accordance with the ethical principles of the Declaration of Helsinki. Informed consent was obtained.

Measures

Demographics

Demographic variables included age (years), sex, race/ethnicity (white, non-white), hourly usage of Internet (daily), and number of spam e-mails received (daily).

Sexual Performance Item

Participants were asked, "Do you believe that you have sexual performance problems?" Response choices were "yes" or "no."

Spam E-mail Items

These items were: 1) Did you receive spam e-mail about sexual performance in the past year?, 2) If yes, did you open and read the e-mail?, and 3) If you opened and read the e-mail, did you purchase anything from the website provided?

Statistical Analyses

Descriptive statistics were calculated. We compared the responses to presence of sexual performance problems to the questions for receiving, opening, and purchasing from spam e-mail on the topics of sexual performance. As relevant, Pearson chi square analyses or the Fisher's exact test were conducted. We conducted a series of three logistic regression analyses for the independent variable of presence of sexual performance problems. The first had the dependent variable of receiving sexual performance spam e-mail, the second had the dependent variable of opening sexual performance spam e-mail, and the third had the dependent variable of purchasing from sexual performance spam e-mail. For each of these three logistic regression analyses, two models were conducted. The first model was an univariate analysis containing just the item inquiring about presence or absence of sexual performance problems. The second model was a multivariate analysis that included the relevant covariates of age, sex, race/ethnicity, hours of Internet use per day, and number of spam e-mails received per day. SPSS version 15 was used for all analyses.

RESULTS

Table 1 describes the demographic characteristics of the sample. The average age was almost 21 years (range: 16 to 28 years). There were more women than men. There was an approximately equal distribution of whites and non-whites. There was on average approximately four hours of Internet use per day (range: 1 to 15 hours). On average there were 28 spam e-mails received per day (range: 0 to 710).

TABLE 1

Demographic Characteristics of a Sample of 200 Individuals and their Spam E-mail Habits

Variable	% (Frequency)	Mean (Standard Deviation)
Age (years)		20.92 (1.99)
Sex		
Men	35.5% (71)	
Women	64.5% (129)	
Race/ethnicity		
White	43.5% (87)	
Non-white	56.5% (113)	
Hourly Internet use per day		3.92 (2.45)
Number of spam e-mails received per day		28.24 (61.86)

Table 2 shows comparisons for sexual performance problems and receiving, opening, and purchasing from spam e-mail regarding sexual performance products. Every individual (100%) with sexual performance problems reported receiving spam e-mail regarding sexual performance products which was significantly different than the approximately 74% from those without sexual performance problems receiving spam e-mail regarding sexual performance products. With regard to opening the spam e-mail for sexual performance products, those who had sexual

performance problems had significantly greater percentages for opening this e-mail than those without sexual performance problems, with a greater than 55% difference. With regard to purchasing the spam e-mail for sexual performance products, those who had sexual performance problems had significantly greater percentages for purchasing from this e-mail than those without sexual performance problems, with a greater than 41% difference.

TABLE 2

Spam e-mail Behaviors and Sexual Performance Problems

Variable	Sexual Performance Problems		p-value
	No % (Frequency) (n=185)	Yes % (Frequency) (n=15)	
Received sexual performance spam e-mail			0.024
No	26.5% (49)	0.0% (0)	
Yes	73.5% (136)	100.0% (15)	
Opened sexual performance spam e-mail			<0.001
No	88.6% (164)	33.3% (5)	
Yes	11.4% (21)	66.7% (10)	
Purchased sexual performance spam e-mail			<0.001
No	94.6% (175)	53.3% (8)	
Yes	5.4% (10)	46.7% (7)	

In the logistic regression analyses, there were no significant differences between those with and without sexual performance problems for receiving spam e-mail for sexual performance products, whether in the univariate or multivariate analyses (data not shown). In the logistic regression analyses for opening spam e-mail for sexual performance products, in the univariate analysis, those with sexual performance problems had significant ($p < 0.001$) odds ratios of 15.62 (95% CI: 4.87, 50.11) than those without sexual performance problems to open spam e-mail for sexual performance products. In the multivariate analysis, a similar significant pattern ($p < 0.001$) occurred with odds ratios of 14.92 (95% CI: 4.60, 48.44). In both the univariate and multivariate analyses none of the covariates were significant.

Table 3 shows the logistic regression analyses for purchasing from spam e-mail for sexual performance products. A similar pattern existed as for opening these spam e-mail messages. In the univariate analysis, those with sexual performance problems had significant odds ratios of greater than 15 times as likely than those without sexual performance problems to purchase from spam e-mail for sexual performance products. In the multivariate analysis, a similar significant pattern occurred with odds ratios of greater than 14 times as likely for those with sexual performance problems to purchase from the spam e-mail for sexual performance products. None of the covariates were significant.

TABLE 3

Logistic Regression Analyses for Variables Associated with Purchasing from Spam E-mail Regarding Sexual Performance Products

Variables	OR (95% CI)	OR (95% CI)
	Model 1	Model 2
Sexual performance problems		
No	1.00	1.00
Yes	15.31 (4.62, 50.74)***	14.34 (4.27, 48.16)***
Age		1.19 (0.91, 1.57)
Sex		
Men		1.00
Women		0.73 (0.24, 2.22)
Race		
White		1.00
Non-white		1.12 (0.37, 3.42)
Hours Internet		1.39 (0.16, 12.31)
Number of Spam e-mails		1.00 (0.99, 1.01)

Note: OR=odds ratio, CI=confidence interval, ***=p<0.001

DISCUSSION

This study suggests that those who have sexual performance problems are highly receptive to spam e-mail for sexual performance products. Even after considering relevant covariates, they are more likely than those without sexual performance problems to open these e-mails and to purchase from these e-mails.

Typically, both men and women find sexually oriented spam as offensive [8]. In our study it appears that those with sexual performance problems are not finding these sexually oriented spam e-mails as offensive while those without sexual performance problems may consider these e-mails offensive. Two-thirds of the individuals with sexual performance problems opened and almost half purchased from this spam e-mail for sexual performance products. On the other hand, among those without sexual performance problems, only a little more than 10% opened and only 5% purchased from the spam e-mail for sexual performance products, suggesting a strong non-interest in these e-mails.

In our logistic regression analyses for both opening and purchasing, personal demographic characteristics and Internet use characteristics were not related to opening or purchasing from spam e-mail for sexual performance products. Consistently for both opening and purchasing from spam e-mail for sexual performance products, there were extremely high odds ratios for those with sexual performance problems to open and purchase from spam e-mail for sexual performance products.

Why is this opening and purchasing behavior from spam e-mail occurring among those with sexual performance problems, especially when the products advertised may not be truly effective

cures? First, these consumers have a perceived medical and/or psychological need to enhance their sexual performance and they are willing to consider products from all sources. Second, this may be a sensitive topic and the private nature of e-mail may be the best venue for purchasing these products rather than interacting face-to-face with a sales clerk at a pharmacy or health food store. Third, there is a perception among individuals who view television drug advertisements that all advertised products are completely safe [7]. This same perception may be occurring among consumers who view drug and other sexual performance product advertisements from spam e-mail.

There are some weaknesses to this study. First, this is not from a national sample. Second, those with sexual performance problems only consisted of 15 individuals. Third, there may be some recall bias where those with sexual performance problems report receiving more spam e-mail.

In conclusion, spam e-mail for sexual performance products are being looked at and purchased from, especially among those who are sensitive to sexual performance concerns. Clearly, those who are sending these spam marketing e-mails have a receptive audience. We are not advocating sending spam e-mail. However, clearly there is a demand for e-mail marketing of sexual performance products and this should be done through ethical and appropriate marketing channels.

REFERENCES

- [1] Fallows, D. "Spam 2007." *Pew Internet & American Life Project*. 2007. Retrieved September 30, 2007 from http://www.pewinternet.org/pdfs/PIP_Spam_May_2007.pdf .
- [2] Fox, S. "Online health search 2006." *Pew Internet & American Life Project*. 2006. Retrieved September 30, 2007 from http://www.pewinternet.org/PPF/r/190/report_display.asp .
- [3] Gernburd, P., Jadad, A. R. "Will spam overwhelm our defenses? Evaluating offerings for drugs and natural health products." *PLoS Medicine*, 2007, 4(9), e274. Retrieved September 30, 2007 from <http://medicine.plosjournals.org/perlserv/?request=get-document&doi=10.1371/journal.pmed.0040274> .
- [4] Grimes, G. A., Hough, M. G., Signorella, M. L. "Email end users and spam: Relations of gender and age group to attitudes and actions." *Computers in Human Behavior*, 2007, 23, 318-332.
- [5] Grimes, G. A. "Online behaviors affected by spam." *Social Science Computer Review*, 2006, 24(4), 507-515.
- [6] Johnson, P.A. "What commercial email contributes to the U.S. economy." New York: Direct Marketing Association, 2003.
- [7] Joseph, M., Stone, G., Haper, J., Stockwell, E., Johnson, K., Huckaby, J. "The effect of manufacturer-to-consumer prescription drug advertisements: an exploratory investigation." *Journal of Medical Marketing*, 2005, 5(3), 233-244.
- [8] Khoo, P. N., Senn, C. Y. "Not wanted in the inbox!: Evaluations of unsolicited and harassing e-mail." *Psychology of Women Quarterly*, 2004, 28, 204-214.
- [9] Mailshell. "Results of the spamcatcher attitude survey." 2003. Retrieved September 30, 2007 from <http://www.mailshell.com/mail/client/oem2.html/step/pr/article/17> .
- [10] Wyke, A. "Patient views: DTC advertising." *International Journal of Medical Marketing*, 2004, 4(4), 310-312.

A CONTINUING STUDY OF IMPACT OF DRESS MODEL TECHNOLOGY ON INTENTION TO BUY ON EVOLVING E-CRM EUROPEAN UNION WEB SITES

James P. Lawler, Pace University, New York City, USA, jlawler@pace.edu, 212-346-1013
Pascale Vandepuette, University of Mons-Hainaut, Belgium, pascale.vandepuette@umh.ac.be

ABSTRACT

This study analyzes design factors that contribute to the intention of goal-focused consumers to buy apparel on European Web sites that have dress model technology. Factors from academic literature of perceived ease of use, perceived usefulness, media richness of technology, decision satisfaction and behavioral intention to return are especially examined in the study. The research-in-progress results begin to confirm earlier exploratory studies by the authors that media richness of dress model technology is less important in intention to buy than the other factors and begin to indicate that the technology is not important in facilitating e-CRM on the Web.

Keywords: business-to-consumer (B2C), dress model technology, e-Business, e-CRM, and European Union

BACKGROUND

Electronic business (e-Business) continues to benefit from the growth in the number of consumers on the Internet, which is currently estimated to be 750 + million globally, a 10% increase from 2006 (Government Technology, March 6, 2007). The number of consumers on the Internet in America in contrast is estimated to be 150 + million, 20% of the 750 + million (Government Technology, March 6, 2007). From a 2006 networked readiness index of 122 countries of the World Economic Forum, countries in Europe are considered to be applying the benefits of information and communication technology (ICT) and of the Web better in competitiveness than in America (O'Connor, 2007). In the index America is exceeded by 5 countries in Europe and is seventh in readiness in 2006 from first in 2005 (Business Week, 2007). Emerging is an ecosystem of innovation in the European Union (Government Technology, March 5, 2007). Enabling benefits of e-Business and the networked readiness of countries in the European Union is the connectivity of the Web.

Networked readiness is enabling a foundation for competitiveness in business-to-consumer (B2C) and customer relationship management (CRM) in countries of the European Union. Finland, France, Germany, Poland and the United Kingdom exceed other European Union countries in indicators of B2C of the European Commission.

Finland, France and the Netherlands exceed in indicators of CRM and customized marketing and sales systems.

Countries in the European Union having high networked readiness furnish benefits for consumers and firms to flourish in a B2C global economy (Carlin, 2007). Estimates indicate

increased competitiveness in B2C and interaction with customers on the Web (e-CRM) and in ICT by firms in the European Union in 2007 – 2011 (Government Technology, March 30, 2007).

This growth of ICT, e-CRM and B2C is contributing to creative design of Web sites that differentiate experiences for consumers in the European Union to buy on the Web. Experiences have to be perceived to be fully functional, engaging and compelling in order to convert consumers into customers (Pullman & Gross, 2004). Such experiences form a foundation for one-to-one relationship marketing and service on the Web that impact the intention of customers to continue to buy on the differentiated sites (Chen, Chen, & Kazman, 2007, pp. 2-3). The inclusion of marketing and self-service is indicated as a number of e-CRM functions or processes (Wagner & Zubey, 2007) that will be improved on European Web sites.

Innovation in the design of buying experiences and relationships and in the integration of technology may give an edge in competitiveness to especially apparel dress model European Union Web sites, if the design and integration on the sites furnish discernable differential for the customers, which is the focus of this study.

INTRODUCTION

Firms in Europe and in America continue to differentiate apparel Web sites by enhancing dress models of their products with media richness (Lawler, Vandeputte, & Joseph, 2007 & Lawler & Joseph, 2006). Literature distinguishes richness as an enabling factor in a mediated environment that affects consumer decisions to buy the products (Klein, 2003; Coyle & Thorson, 2001; & Bezjian-Avery & Calder, 1998) and be loyal to the sites (Gharbi & Soltani, 2007). Richness is defined as multimedia representations of products for enhancing not only the pleasure of shopping but also the process of shopping on these sites and simulating the senses of shoppers (Klein, 2003; Coyle & Thorson, 2001; & Bezjian-Avery & Calder, 1998). Richness is evident in consumer customized 3D dynamic imaging of products on models improved with multimedia technology. Such models evolved from impersonal default 2D mannequins (Greenemeier, 2005) on apparel Web sites. Interactivity of the models may affect experiences to buy the products (Fiore & Jin, 2003; Li, Daugherty, & Biocca, 2001). Interactivity may be perceived to be more realistic (Hoffman & Novak, 1996; Novak, Hoffman, & Yung, 1999) and potentially recreational and social (Hof, 2007) on the sites than interaction in the physical stores. Investment in media rich systems, improved by broadband and digital imaging that integrate Asynchronous Javascript and XML (AJAX) (Danziger, 2006) and Flash on product screens, P2P streaming of product shows (Schonfeld, 2007), and potentially Entropia Universe, Second Life and There simulations (McConnon, 2007; Newitz, 2007; & Hemp, June, 2006) of stores on vivid 3D Net Web sites, is justified in the perception that richness of technology is important in intention to buy on apparel Web sites if not important in loyalty to the sites.

Research into the media richness of dress model technology is documented in exploratory studies of European and American apparel Web sites by the authors of this study in 2007 (Lawler, Vandeputte, & Joseph, 2007) and 2006 (Lawler & Joseph, 2006). Results of the studies indicate that the factor of richness of technology is less important than non-technological factors in the intention of goal-focused knowledgeable customers to buy on apparel Web sites. Factors of control and freedom, product selection, lack of sociality, availability of information and

accessibility and convenience identified in the studies from a narrow niche of the practitioner literature (Wolfenbarger & Gilly, 2001) indicate more importance than the hype of richness of technology in the intention of goal-focused customers to buy products on Web sites. Practitioner studies further indicate the importance of non-technological factors (Nielsen, 2007) on Web 1.0 sites, and the non-importance of Web 2.0 technological trends, as consumers may form flash impressions of sites based frequently on non-technological factors (McCoy, Everard, Polak, & Galletta, 2007) and surf to other sites on which they buy the products (Srinivasan, Anderson, & Ponnnavolu, 2002). These studies form a foundation for designing evolving e-CRM functionality that integrates media richness of technology on European Union apparel sites, so that consumers buy and customers continue to buy on these sites.

This study extends the factors of control and freedom, product selection, lack of sociality, availability of information and convenience and accessibility, from the narrow dimension of generally practitioner literature of the studies in 2007 (Lawler, Vandeputte, & Joseph, 2007) and 2006 (Lawler & Joseph, 2006). Factors of perceived ease of use, perceived usefulness, decision satisfaction and behavioral intention to return, from primarily academic literature, furnish a broader dimension for evaluating the relative importance of richness. They furnish a foundation for further evaluating the feature functionalities of e-CRM systems in importance in the initiation of customer relationship on European Union dress model sites. This study helps e-Business managers in the European Union as firms investigate future technologies for interactivity on the Web (Orman, 2007) that may impact and be important in intention to buy and to continue to buy on apparel dress model sites. This study may be helpful in improving return-on-investment (ROI) (Pratt, 2007) on sites that intelligently integrate richness of media technology and e-CRM.

FACTORS

The factors of this study are summarized below in Table 1:

Table 1: Factors of Perception in Intention to Buy of Goal-Focused Shoppers on European Web

Factor	Definition	Primary Sources
Pre-Attitude to Web Shopping	Extent of Favorable Preference to Buy Apparel Products on Dress Model Sites of Web	Cheung, Chan, & Limayem, 2005
Perceived Ease of Use of Web Site	Extent of Perception of Freedom in Effort and Navigation on Dress Model Sites	Davis, 1989
Perceived Usefulness of Web Site	Extent of Perception of Functionality in Enhancing Process of Apparel Shopping on Dress Model Sites	Davis, 1989
Richness of Media Technology on Web Site	Extent of Perception of Multimedia Representations of Products on Dress Model and Picture Sites in Enhancing Process of Apparel Shopping on Sites	Lawler, Vandeputte, & Joseph, 2007; Lawler & Joseph, 2006; Klein, 2003; Coyle & Thorson, 2001; & Bezjian-Avery & Calder, 1998
Decision Satisfaction with Web Site	Extent of Perception of Support and Tools on Dress Model Sites in Enhancing Process of Apparel	Garrity, Glassberg, Kim, Sanders, & Shin, 2005

	Shopping on Sites	
Behavioral Intention to Return to Web Site	Extent of Favorable Perception of Dress Model Sites in Enhancing and Using Information and Process of Apparel Shopping on Sites	Venkatesh, Morris, Davis, & Davis, 2003
Post-Attitude to Web Shopping	Extent of Favorable Preference to Continue to Buy Apparel Products on Dress Model Sites of Web	Cheung, Chan, & Limayem, 2005

FOCUS

The focus of this study is to evaluate if not finalize the importance or non-importance of richness of media technology as a design factor relative to the mix of non-technological factors of perceived ease of use, perceived usefulness, decision satisfaction and behavioral intention to return, in the intention of European goal-focused customers to buy and to continue to buy products on European Union dress model and picture sites. Fresh insight into the non-importance or importance of richness of media technology may be helpful in informing managers of improved design strategy on the Web, at a time when the European Union is expanding funding for research in ICT (Blau, 2007 & Government Technology, November 21, 2006) and next generation multimedia technologies (Gardner, 2006). This study is timely.

METHODOLOGY

The experiment of this study continues from a base of analysis that began as a project in Strategies for e-Commerce Technologies course, at Pace University in America in 2003, and continued as projects in an e-Business Concepts and Applications course, at the University of Mons-Hainaut in Belgium into 2007.

The methodology of the current analysis consisted of a sample of 18 European apparel Web sites chosen by 36 primarily Belgian students in the spring 2007 semester. They were chosen by the authors from the frequency of e-Business and e-CRM functionalities on the sites. The bulk of the sites were also the focus of analysis by the authors and previous students in the 2007 (Lawler, Vandeputte, & Joseph, 2007) study. The sites were analyzed anonymously and independently by the students in a Bell Net Ethernet Microsoft NT broadband laboratory in a 5 month period of the spring semester as an extra credit program of the course. These students were proactive and realistic Web wired subjects (Ohmae, 2005) for the experiment, as in the previous study.

The 18 sites in the 2007 study were evaluated by the 36 students for design factors in intention to buy apparel on the sites. Each student was allocated a fictitious 1,500 euros to buy products from any number of the 18 sites for a fictitious goal-focused task of being casually dressed for a forthcoming summer vacation in Tunisia. They analyzed each of the sites on the factors of perceived ease of use of the Web site, perceived usefulness of the site, richness of media technology on the site, decision satisfaction with the site, and behavioral intention to return to the site. They also analyzed pre-attitude to Web shopping and post-attitude to Web shopping from their experience of the experiment. They assigned a fine 7-point scale of 7 = very high in importance in intention to buy, 6 = high in importance, 5 = somewhat high in importance, 4 = somewhat low in importance, 3 = low in importance, 2 = very low in importance, and 1 = not at

all important in the project shopping task, even if they did not buy on a site. The experiment was done with instructions and an instrument in the French language, which were evaluated for feasibility in a pilot for the shopping task by 3 instructors and 6 students of the university. The 36 students finished the task in July of the spring semester and furnished the data for analysis to the authors of this study.

PRELIMINARY DATA ANALYSIS

The preliminary analysis of the descriptive data on the 7 factors, 18 sites and 36 students of the current study appears to be indicating the factor of richness of media technology to be lower in importance than perceived ease of use, perceived usefulness, decision satisfaction and behavioral intention to return, in intention to buy on the dress model Web sites. The means of the factors of behavioral intention, decision satisfaction, usefulness and ease of use appear to be higher in importance in the intention of most of the female and male students to buy on most of the sites. The factor of richness of technology appears to have the lowest of means of the sites and of the students, though several of the sites are consistently and demonstratively higher in the means of richness than the other factors, in contrast to the means of richness in the 2007 (Lawler, Vandeputte, & Joseph, 2007) European study. The data from the academic factors in this study is beginning to confirm results on relative lowness of richness in contrast to the primarily practitioner factors of the 2007 and 2006 (Lawler & Joseph, 2006) studies, despite a few improvements in the integration of multimedia technologies on several of the sites.

PRELIMINARY IMPLICATIONS

From a preliminary analysis of the data in the current study, the basic design factors, of perceived ease of use, perceived usefulness, decision satisfaction and behavioral intention to return, continue to be more important relative to the factor of richness of media technology in intention to buy on the European dress model sites. Features of e-CRM functionalities on the sites were also important in the intention to buy and to continue to buy on the sites. Perceived ease of use and perceived usefulness were more important in positive post-attitude by the students to Web shopping. Given the hype of media richness in practitioner literature, preliminary results of this study imply that e-Business managers of B2C dress model sites may have to evaluate the fruitfulness of investment in media technologies in shopping on the Web. This implication is important in the initiation of managers to integrate multimedia technologies into 3D virtual shopping of Web 2.0 (Hemp, October, 2006).

Though the factor of richness of media technology continued to be perceived less important than the fundamental factors of ease of use, usefulness, decision satisfaction and behavioral intention to return, in intention to buy and to continue to buy on the dress model sites in this study, richness was still important in the mix of the non-technological factors. Richness was perceived to be important in shopping on the sites instead of in the stores, but it was not perceived to be optimal in the mix of decision satisfaction and behavioral intention to return nor in the likelihood of relationship. Richness of technology was essentially not perceived as positive as the other factors in influencing shopping. This implication is important in the need to optimally plan

integration of multimedia technologies not only into Web 2.0 sites but into purely Web 1.0 sites of today.

Throughout the preliminary analysis of this study, the bulk of the dress model sites, in contrast to the sites in the 2007 (Lawler, Vandeputte, & Joseph, 2007) study, were not improved noticeably in the integration of media technology in the intention to buy on the European sites. The positive post-attitude to Web shopping was not influenced by richness of media technology on the bulk of the sites. Several sites improved by media technology were however influencing importance in positive post-attitude to Web shopping, by the integration of the multimedia into the mix of the non-technological design of these sites. Post-attitude to shopping on the Web, in contrast to pre-attitude to Web shopping, was nevertheless not improved noticeably overall in the perception of the students in this study. These implications are important in planning prudent investment in multimedia technologies on the sites, in order to satisfy the sensitivity of shoppers in the shopping task.

REFERENCES

References are available upon request of James P. Lawler.

LINKING KNOWLEDGE PROPERTIES AND KNOWLEDGE NETWORK TOPOLOGY WITH PERFORMANCE

Miguel A. Bustamante, Metro State College of Denver, (303) 556-2932, mbustama@mscd.edu
David L. Olson, University of Nebraska, (402) 472-3915, dolson3@unl.edu

ABSTRACT

This research argues that a company's capacity to improve its performance will depend on sound management of its available knowledge. Consequently, this study develops a research model (simulation) that links the properties of knowledge generated by the system and the features of a firm's knowledge network topology to predict performance. First, this research represents knowledge characteristics of the organization. Second, this study expresses the peculiar features of the firm's knowledge network topology. Finally, these concepts are incorporated into a new model to predict the impact that this relationship will have in the firm's ability to solve problems (performance).

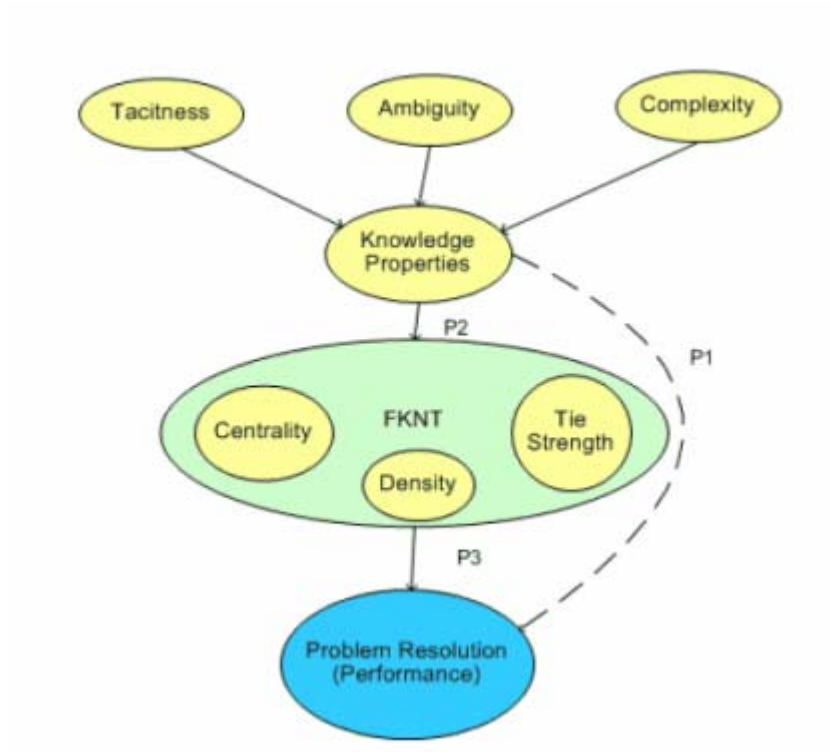
Keywords: Knowledge management, Garbage-can model, Simulation

INTRODUCTION

Information and Communication Technologies (ICT) have well known, but often unfulfilled, potential to alter businesses and industries (Andal-Ancion et al., 2003). Companies all around the world are seeking to improve firm performance by buying more ICT. Yet they find that investing in ICT usually does not generate expected perceived benefits (Tippins and Sohi, 2003). Recent information systems research has identified the main cause of performance gains to be the optimal flow of knowledge generated by the system (Alavi and Leidner, 2001; Grover and Davenport, 2001; Holsapple and Joshi, 2002; Sussman and Siegal, 2003). Furthermore, empirical research shows that in many cases the optimal flow of knowledge generated by the system is achieved through a lengthy and costly process of trial-and-error (Ko et al., 2005). Unfortunately, this approach can misuse precious time and assets, especially in small and medium size U.S. firms where organizations may have limited resources. A methodology that links the properties of knowledge generated by the system and the features of a firm's knowledge network topology may predict performance, and therefore help organizations quickly reach the optimal flow of knowledge.

This paper seeks to address the following question: what function does the relationship between knowledge properties -- tacitness, ambiguity, and complexity -- (Nonaka, 1994; Szulanski, 1996; Hansen, 1999; Alavi and Leidner, 2001; Courtney, 2001; Birkinshaw et al., 2002; Argote et al., 2003; Lin et al., 2005; Nissen, 2006) and the organization's knowledge network topology -- centrality, density, tie strength -- (Hansen, 1999; Reagans and McEvily, 2003; Kane and Alavi, 2005) serve in improving firm performance? In essence, this study deals with all the main properties of knowledge and seeks to understand the relationship of all these properties with the main characteristics of knowledge network topology.

OPERATIONAL RESEARCH MODEL



RESEARCH MODEL DESCRIPTION

The model begins by presenting a superset (generic set) of the main properties of knowledge (tacitness, ambiguity, and complexity) it implies that combinations of the different degrees of knowledge's properties will indirectly impact the organization's problem resolution abilities (performance). Knowledge's impact is mediated by the firm's network topology. The model then operationalizes the argument that knowledge's impact is mediated by the firm's network topology critical features: tie strength, density, and centrality (Hansen, 1999; Reagans and McEvily, 2003; Kane and Alavi, 2005). Tie strength effects are tested by comparing the frequency with which individuals interact with the system. Density effects are tested by comparing the ratio of actual ties to the number of possible ties in the network. Centrality effects are tested by comparing the number of individuals who act as a knowledge broker for the rest of the network (Hansen, 1999; Te'eni, 2001; Griffith et al., 2003; Reagans and McEvily, 2003; Kane and Alavi, 2005; Paul, 2006).

Model's propositions

The model formulates three propositions:

Proposition 1: A firm's knowledge network topology will mediate the impact that knowledge has on the organization's problem resolution abilities (performance).

Proposition 2: Organizational performance is sensitive to variations in knowledge properties.

Proposition 3: When the firms' knowledge network topology matches with the particular property of the knowledge that is going to be communicated, the organization's problem resolution abilities (performance) will be optimal.

Testing propositions

To test these propositions this study will use a garbage can (GC) simulation model (Cohen et al., 1972). The primary difference between the canonical GC model and our model is that Cohen et al. (1972) fixed the net energy load of the organization to three levels in the canonical GC model. In our model, the energy load of the organization is linked to knowledge properties. For example, when tacitness is low, ambiguity is low, and complexity is low, the energy required to solve a problem will be at a minimum. Conversely, if all three knowledge properties are high, energy required to solve a problem will be at a maximum. The main constructs of the GC model (access structure, decision structure, and energy distribution) are matched to the firm's knowledge network topology characteristics (tie strength, centrality, and density). Based on the interrelationships of the firm's knowledge properties and the firm's knowledge network topology characteristics while considering the number of problems fixed to 20 and the number of decision makers fixed to 10 the simulation model determines three dependent variables. First, the number of choices solved; second, the level of problem activity; and finally, the level of decision difficulty.

EXPERIMENTAL RESULTS

The model was used to examine the impact of network topology on system performance. Three factors were varied. First, Centrality represents the role of information systems in the network, (individuals interact with the system directly or they need intermediaries to interact with the system). Second, Tie Strength represents how people do their work. High tie strength represents environments with high levels of interactive use of an information system. Low tie strength would involve batch submission of requests of the information system. Mixed would involve both. Finally; Density represents the number

of relationships in the network relative to the possible number. High density relationships would have all network members interacting. Low density would have very limited interactions among pairs of participants, while mixed was an intermediate level. Outcome measures by different levels of centrality, tie strength, and density are shown in Table 1, Table 2, and Table 3 respectively

TABLE 1: SIMULATION RESULTS BY CENTRALITY

Measure	Centrality	C00	C01	C02	C10	C11	C12	C20	C21	C22
Solved Choices	Independent	5.41	5.51	5.33	4.89	4.46	3.79	3.22	3.04	2.89
	Interdependent	3.94	3.76	3.57	3.22	2.78	2.51	2.11	1.86	2.03
	Dependent	2.11	1.92	1.83	1.56	1.35	1.05	0.69	0.63	0.48
Decision Difficulty	Independent	10.87	13.74	17.08	21.39	23.54	25.55	27.50	30.44	32.02
	Interdependent	22.17	26.85	29.33	32.19	33.45	36.45	36.40	36.54	37.58
	Dependent	40.53	41.71	42.49	45.31	47.76	49.94	51.54	52.11	52.78
Problem Activity	Independent	57.47	65.52	85.02	115.19	143.51	169.80	189.36	196.98	204.33
	Interdependent	127.02	138.65	153.62	168.83	188.44	195.14	208.78	219.59	215.34
	Dependent	229.56	237.42	243.25	248.56	252.45	256.06	258.34	259.69	261.03

TABLE 2: SIMULATION RESULTS BY TIE STRENGTH

Measure	Tie Strength	C00	C01	C02	C10	C11	C12	C20	C21	C22
Solved Choices	High	2.30	2.44	2.28	1.78	1.41	1.02	0.39	0.19	0.28
	Mixed	3.28	3.04	2.84	2.59	2.16	1.72	1.42	1.19	1.18
	Low	5.88	5.70	5.62	5.31	5.03	4.60	4.22	4.16	3.94
Decision Difficulty	High	20.46	22.43	20.70	20.31	19.87	21.44	20.89	21.06	21.15
	Mixed	25.35	27.85	30.60	35.30	37.71	39.39	40.45	42.29	43.63
	Low	27.75	32.03	37.60	43.28	47.17	51.11	54.09	55.75	57.60
Problem Activity	High	208.07	209.63	221.30	238.11	257.78	274.15	294.22	301.59	300.41
	Mixed	153.91	171.52	189.38	212.11	236.78	249.40	258.97	268.09	270.05
	Low	52.07	60.44	71.21	82.37	89.85	97.45	103.29	106.58	110.25

TABLE 3: SIMULATION RESULTS BY TIE DENSITY

Measure	Density	C00	C01	C02	C10	C11	C12	C20	C21	C22
Solved Choices	Low	3.33	3.39	3.28	2.82	2.58	2.26	1.99	1.81	1.73
	Mixed	3.97	3.76	3.75	3.51	3.06	2.37	1.81	1.69	1.66
	High	4.15	4.04	3.71	3.34	2.95	2.71	2.22	2.03	2.01
Decision Difficulty	Low	27.59	30.09	31.74	34.18	35.68	37.54	38.15	39.41	40.07
	Mixed	23.41	27.36	29.87	33.59	35.92	38.35	39.64	41.08	41.94
	High	22.56	24.85	27.30	31.12	33.16	36.06	37.65	38.60	40.37
Problem Activity	Low	149.45	153.74	164.53	184.29	200.18	207.80	218.83	226.78	228.28
	Mixed	137.41	149.06	160.84	175.30	192.44	211.16	221.15	224.76	228.03
	High	127.19	138.78	156.52	173.01	191.79	202.05	216.50	224.73	224.40

CONCLUSIONS

This study analyzed the relationship between knowledge properties and organizational knowledge network topology with respect to knowledge transfer efficiency. The garbage can model of Cohen et al. (1972) was used as a basis for development of a simulation model to calculate expected performance given a complex set of inputs, focusing on how different combinations of energy required to make decisions relates to different network topology configurations. In model calibration, solved choices varied from 50 percent to 9 percent as decision making energy requirements were increased, decision difficulty from 80 percent to 9 percent, and problem activity from 80 percent to 10 percent.

These results are consistent with prior research (Hansen, 1999; Argote et al., 2003; Lin et al., 2005; and Nissen, 2006) in that organizational performance is sensitive to variations in firm knowledge properties. The study extends the research of Reagans and McEvily (2003), demonstrating that various properties of knowledge can be separated. It also extends the research of Uzzi and Lancaster (2003) by demonstrating how network topology can be modeled in terms of centrality, tie strength, and density.

Density Conclusions

Findings with respect to density are that solved choices decreases as energy requirements increase for all three types of density studied. Therefore Proposition 2 (Organizational performance is sensitive to variations in knowledge properties) is supported. The proportion of solved choices was found to increase for low density networks as energy required increased. Mixed network density was found to have a transitional relative advantage with respect to proportion of choices solved. As energy required increased, decision difficulty and problem activity per model were found to decrease for low density networks. Reagans and McEvily (2003) and Kane and Alavi (2005) found that high network density

would have higher performance. Our study conversely indicates that low network density may have relative advantage when level of energy required increases.

Tie Strength Conclusions

As decision making energy required increased solved choices decreased for all three types of tie strength modeled, Therefore Proposition 1 (A firm's knowledge network topology will mediate the impact that knowledge has on the organization's problem resolution abilities) is supported.. When tie strength was low, the proportion of solved choices increased with increasing energy required, while when tie strength was mixed or high, the proportion of solved choices decreased. This extends Reagans and McEvily (2003) by exploring different types of tie strength under varying levels of decision making energy required. Our results indicate the robustness of low tie strength as energy required increases. A firm's knowledge network topology will mediate the impact that knowledge has on the organization's problem resolution ability. The proportion of decision difficulty and problem activity increased for low and mixed tie strengths, but for high levels of tie strength the proportional level of decision difficulty decreased while problem activity remained high. A plausible explanation may be that when the type of tie strength is inappropriate for a given circumstance, problem activity may be high, but few decisions are made, and thus the number of solved choices decreases.

Centrality Conclusions

The number of solved choices decreased as decision making energy required increased for all types of centrality. Under independent centrality, the proportion of solved choices were found to increase with increasing energy requirements, while dependent centrality involved a decrease in solved choices per model. A mix of dependent and interdependent centrality was found to perform better than independent centrality for low levels of decision making energy required. Kane and Alavi (2005) studies network setups versus performance without varying energy required for problem solution. They concluded that dependent centrality would have low performance relative to independent centrality. Our findings do not refute that conclusion, but infer that a combination of dependent and interdependent centrality might outperform independent centrality for low levels of decision making energy required. Therefore Proposition 3 (When the firms' knowledge network topology matches with the particular property of the knowledge that is going to be communicated, the organization's problem resolution abilities (performance) will be optimal) is supported.

Final Conclusions

This research demonstrates an approach to building theory considering the contribution of social network theory and the sociology of knowledge (Hansen, 1999; Kane and Alavi, 2005) in conjunction with behavioral assumptions (Cohen et al., 1972). It combines these theories with information systems theory (Reagans and McEvily, 2003). It raises the issue of knowledge properties (tacitness, ambiguity, complexity) and their impact on information transfer (Uzzi and Lancaster, 2003) with respect to different elements of knowledge network topology (centrality, density, tie strength).

REFERENCES

References available upon request from Miguel A. Bustamante, Metro State College of Denver, (303) 556-2932, mbustama@mscd.edu

OPTIMAL KNOWLEDGE REFRESHING POLICES

Xiao Fang

Department of Information Operations and Technology Management, University of Toledo, Toledo, OH 43606, xiao.fang@utoledo.edu

ABSTRACT

Knowledge Discovery in Databases (KDD) gives organizations the tools to sift through vast data stores to extract knowledge supporting organizational decision making. In this paper, we research knowledge refreshing, which we define as the process to keep knowledge discovered using KDD up-to-date with its dynamic data source. We propose a Markov decision process model for knowledge refreshing and derive optimal knowledge refreshing policies from the model.

INTRODUCTION

Knowledge Discovery in Databases (KDD) [2] gives organizations the tools to sift through vast data stores to extract knowledge supporting organizational decision making. Most of the KDD research has assumed that data is static and focused on either efficiency improvement of the KDD process or business applications of KDD. However, data is dynamic in reality (i.e., new data continuously added in). Knowledge discovered using KDD becomes obsolete over time, as the discovered knowledge only reflects the status of its dynamic data source when running KDD. Newly added data could bring in new knowledge and invalidate some discovered knowledge. To support effective and efficient decision making, knowledge discovered using KDD needs to be updated along with its dynamic data source. In this research, we focus on knowledge refreshing, which we define as the process to keep knowledge discovered using KDD up-to-date with its dynamic data source.

Related research, which investigated how to maintain patterns learned from data mining over a dynamic data source, includes: incremental data mining [1] and data stream mining [4]. Both incremental data mining and data stream mining research only addressed one step in the KDD process – the data mining step. However, to support effective decision making, the KDD process needs to be completed. The KDD process is a highly costly process [5]. As a result, it is impractical to run KDD whenever there is an update in a data source. It is also unnecessary to run KDD whenever there is an update in a data source. Such a practice often results in repetitive knowledge identical with previous KDD run because successive snapshots of real world data sources overlap considerably [3]. On the other hand, running KDD too seldom could result in losing critical knowledge, which, in turn, will impact decision making negatively. Therefore, it is critical to determine when to run KDD to optimize the trade-off between the cost of knowledge loss and the cost of running KDD.

MODEL

As shown in Figure 1, a data source evolves over time as new data continuously populated into it. To keep the knowledge base up-to-date with the data source, a KDD run has to be executed. There are queries submitted to the knowledge base to get knowledge supporting organizational decision making. At the time of a query arrival, executing KDD incurs cost of running KDD while not executing KDD results in cost of knowledge loss. The knowledge refreshing problem addressed in this research is to find rules determining time points to execute KDD so that the total cost, including costs of running KDD and costs of knowledge loss, over the time horizon is minimized.

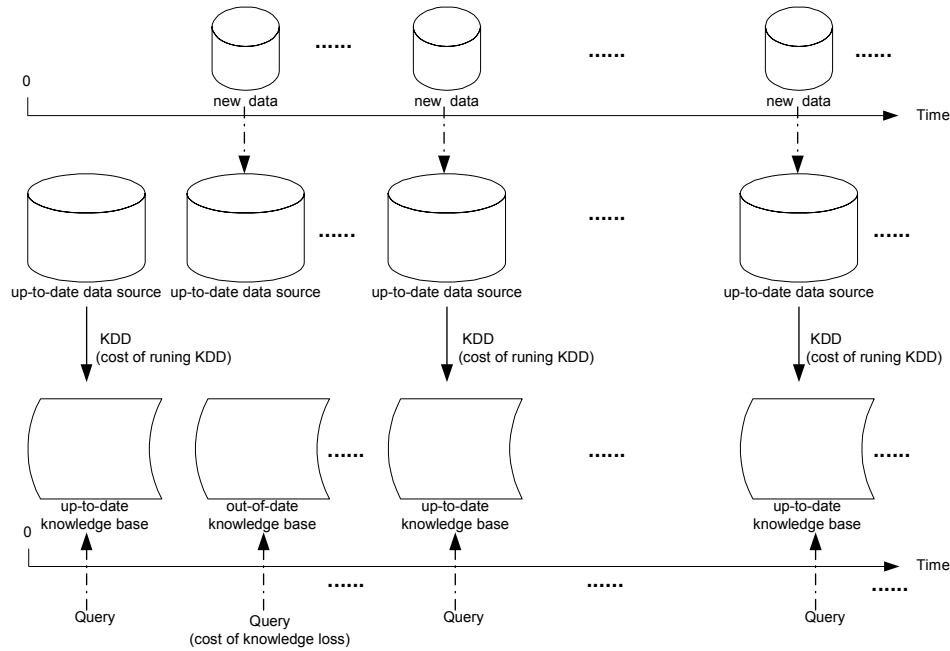


Figure 1: Knowledge Refreshing

The knowledge refreshing problem is modeled as a Markov decision process. We assume that new data arrival follows a Poisson process with intensity λ_d . Queries for the knowledge base are classified into distinct types based on the following consideration. Different types of queries have different arrival rates and the same amount of knowledge loss has different impact on different types of queries. For example, queries supporting strategic decision making could have lower arrival rate than queries supporting operational decision making; while the same amount of knowledge loss could have higher impact on the former than the latter. We assume that there exists n distinct types of queries, $\{q_i\}, i = 1, 2, \dots, n$, and each type of queries follows a Poisson process with intensity λ_{q_i} independently. Query arrivals are also independent of new data arrival. The decision points are the moments when a query for the knowledge base arrives. The system time of the model is labeled as $t = 0, 1, 2, \dots, m, \dots$, with 0 being the starting time of the time horizon and m , where $m \geq 1$, denoting the time of decision point m or the time when the arrival of the m th query. The action space of the model is defined to be a binary set $A = \{0, 1\}$, where 0 denotes not running KDD and 1 denotes running KDD. The system state at decision point m is represented by a vector $s_m = (q_m^a, l_m)$, where q_m^a denotes the type of the query arrived at decision point m and l_m denotes the amount of knowledge loss at decision point m . According to the independent Poisson arrival assumption for queries, the probability, $P\{q_m^a = q_i\}$, that type i query arrives at decision point m , is $\lambda_{q_i} / \sum_{j=1}^n \lambda_{q_j}$. Let r_m be the latest decision point with action 1 (i.e., running KDD) before decision point m . l_m is determined by d_m , the number of units (e.g., blocks) of new data arrived between decision points r_m and m , and l , the expected amount of knowledge loss brought in by a unit of new data.

$$l_m = d_m l \quad (1)$$

Substituting d_m in (1) with its expected value $\lambda_d (m - r_m) / \sum_{j=1}^n \lambda_{q_j}$, we have,

$$l_m = (m - r_m) \lambda_d l / \sum_{j=1}^n \lambda_{q_j} \quad (2)$$

According to (2), the set of all possible l_m at decision point m is L_m , where $L_m = \{\lambda_d l / \sum_{j=1}^n \lambda_{q_j}, 2\lambda_d l / \sum_{j=1}^n \lambda_{q_j}, \dots, m \lambda_d l / \sum_{j=1}^n \lambda_{q_j}\}$. In summary, the state space S_m at decision point m can be defined as:

$$S_m = \{s_m | s_m = (q_m^a, l_m), \text{ where } q_m^a \in \{q_i\} \text{ and } l_m \in L_m\} \quad (3)$$

Let $a_m \in A$ be the action chosen at decision point m . One step transition from state $s_m = (q_m^a, l_m)$ at decision point m to state $s_{m+1} = (q_{m+1}^a, l_{m+1})$ at decision point $(m+1)$ under action a_m consists of two stages. The first stage is the transition from l_m to l_{m+1} . Given l_m and a_m , there is only one possible amount of knowledge loss at next decision point, l_{a_m} , where,

$$l_{a_m} = \begin{cases} l_m + \lambda_d l / \sum_{j=1}^n \lambda_{q_j} & \text{if } a_m = 0 \\ \lambda_d l / \sum_{j=1}^n \lambda_{q_j} & \text{if } a_m = 1 \end{cases} \quad (4)$$

Thus, the transition probability $P_{l_m l_{m+1}}^{a_m}$ from l_m to l_{m+1} under action a_m is,

$$P_{l_m l_{m+1}}^{a_m} = \begin{cases} 1 & \text{if } l_m = l_{a_m} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

The second stage is the transition from q_m^a to q_{m+1}^a , which results from the arrival of query q_{m+1}^a . Hence, the transition probability of the second stage is $P\{q_{m+1}^a = q_i\}$, where $i = 1, 2, \dots, n$. The second stage transition is independent of the first stage transition as query arrival is independent of new data arrival and the action chosen. As a result, one step transition probability $P_{s_m s_{m+1}}^{a_m}$ from state s_m to s_{m+1} under action a_m is determined as the product of $P_{l_m l_{m+1}}^{a_m}$ and $P\{q_{m+1}^a = q_i\}$,

$$P_{s_m s_{m+1}}^{a_m} = \begin{cases} P\{q_{m+1}^a = q_i\} & \text{if } l_m = l_{a_m} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where $P\{q_{m+1}^a = q_i\} = \lambda_{q_i} / \sum_{j=1}^n \lambda_{q_j}$. The system cost $c(s_m, a_m)$ incurred at decision point m is,

$$c(s_m, a_m) = \begin{cases} c_m^l & \text{if } a_m = 0 \\ c_m^k & \text{if } a_m = 1 \end{cases} \quad (7)$$

where c_m^l denotes the cost of knowledge loss at decision point m and c_m^k denotes the cost of running KDD at decision point m . For different types of queries in $\{q_i\}, i = 1, 2, \dots, n$, let the cost of per unit knowledge loss be c_{q_i} respectively.

$$c_m^l = c_{q_i} l_m \quad \text{if } q_m^a = q_i, i = 1, 2, \dots, n \quad (8)$$

Substituting l_m in (8) with (2), we have,

$$c_m^l = (m - r_m) \lambda_d l c_{q_i} / \sum_{j=1}^n \lambda_{q_j} \quad \text{if } q_m^a = q_i, i = 1, 2, \dots, n \quad (9)$$

Although the computational cost of data mining is proportional to the size of the mined data, the cost of data pre-processing and pattern post-processing, which usually involves people, typically does not vary with data volume and the latter dominates in the cost of running KDD. Hence, we model the expected cost c_m^k of running KDD at decision point m as a constant C_k . Therefore, the system cost $c(s_m, a_m)$ incurred at decision point m becomes,

$$c(s_m, a_m) = \begin{cases} (m - r_m) \lambda_d l c_{q_i} / \sum_{j=1}^n \lambda_{q_j} & \text{if } a_m = 0 \text{ and } q_m^a = q_i, i = 1, 2, \dots, n \\ C_k & \text{if } a_m = 1 \end{cases} \quad (10)$$

The expected number of decision points over a relative long time horizon, M , is $T \sum_{j=1}^n \lambda_{q_j}$, where T is the duration of the time horizon. A refreshing policy π can be defined as $\pi = (\delta_1, \delta_2, \dots, \delta_m, \dots, \delta_M)$, where δ_m is the decision rule at decision point m such that $a_m = \delta_m(s_m)$. The expected total system cost over the time horizon under refreshing policy π , EC_π , can be expressed as,

$$EC_\pi = E \left\{ \sum_{i=1}^M [\delta_i(s_i) c_i^k + (1 - \delta_i(s_i)) c_i^l] \right\} \quad (11)$$

The optimal knowledge refreshing policy, π^* , minimizes expected total system cost over the time horizon such that $EC_{\pi^*} = \min_{\pi \in \Pi} EC_\pi$, where policy space Π is the set of all feasible knowledge refreshing policies.

OPTIMAL POLICIES

A dynamic programming solution recursively evaluates the individual knowledge refreshing decisions over a time horizon in their reverse time order. Let $v_{s_m}(m)$ be the optimal expected total system cost from decision point m to the end of the time horizon with current state s_m , where $1 \leq m \leq M$. $v_{s_m}(m)$ can be calculated as,

$$v_{s_m}(m) = \min_{a_m \in A} \{ c(s_m, a_m) + \sum_{s_{m+1} \in S_{m+1}} P_{s_m s_{m+1}}^{a_m} v_{s_{m+1}}(m+1) \} \quad (12)$$

Let $v(0)$ be the optimal expected total system cost from the start to the end of the time horizon.

$$v(0) = \sum_{s_1 \in S_1} P_{s_1} v_{s_1}(1) \quad (13)$$

In (13), $S_1 = \{s_1 | s_1 = (q_1^a, l_1)\}$, where $q_1^a \in \{q_i\}$ and $l_1 = \lambda_d l / \sum_{j=1}^n \lambda_{q_j}$ and

$$P_{s_1} = \lambda_{q_i} / \sum_{j=1}^n \lambda_{q_j} \quad \text{if } s_1 = (q_i, l_1), \text{ for } i = 1, 2, \dots, n$$

Applying the dynamic programming solution, $v(0)$ can be solved backwards from decision point M . At the same time, EC_{π^*} and π^* are obtained, where $EC_{\pi^*} = v(0)$.

Although tractable, the dynamic programming solution is still computational expensive for large M . A more computationally efficient solution can be obtained by reducing police space from Π to Π_f . Π_f contains only policies refreshing knowledge every fixed number of queries/decision points, namely fixed knowledge refreshing policy space. Motivated by research on determining optimal fixed time intervals between checkpoints in Database Management Systems[5], we introduce Lemma 1 to

determine the fixed optimal knowledge refreshing policy, π_f^* , such that $EC_{\pi_f^*} = \min_{\pi_f \in \Pi_f} EC_{\pi_f}$, where π_f is a fixed knowledge refreshing policy.

Lemma 1. The fixed optimal knowledge refreshing policy π_f^* is a policy that runs KDD every

Q^* queries/decision points (i.e., runs KDD at decision points $Q^*, 2Q^*, \dots$) where $Q^* = \sqrt{\frac{2C_k \sum_{j=1}^n \lambda_{q_j}}{l \bar{c}_q \lambda_d}}$

and $\bar{c}_q = \frac{\sum_{j=1}^n \lambda_{q_j} c_{q_j}}{\sum_{j=1}^n \lambda_{q_j}}$.

Proof. Following a knowledge refreshing policy in Π_f , KDD is run every Q decision points. The expected cost of running KDD over the time horizon, C_{KDD} , is

$$C_{KDD} = C_k \frac{M}{Q} \quad (14)$$

The time horizon can be divided into $\frac{M}{Q}$ periods with each period including Q queries. The expected cost of knowledge loss at each period, $C_{KL/P}$, is

$$C_{KL/P} = \frac{\lambda_d}{\sum_{j=1}^n \lambda_{q_j}} l \frac{Q(Q-1)}{2} \bar{c}_q \quad \text{where } \bar{c}_q = \frac{\sum_{j=1}^n \lambda_{q_j} c_{q_j}}{\sum_{j=1}^n \lambda_{q_j}}.$$

The expected cost of knowledge loss over the time horizon, C_{KL} , is

$$C_{KL} = C_{KL/P} \frac{M}{Q} = \frac{\lambda_d}{\sum_{j=1}^n \lambda_{q_j}} l \frac{M(Q-1)}{2} \bar{c}_q \quad (15)$$

Adding (15) and (16), the expected system cost over the time horizon, EC , is

$$EC = C_k \frac{M}{Q} + \frac{\lambda_d}{\sum_{j=1}^n \lambda_{q_j}} l \frac{M(Q-1)}{2} \bar{c}_q \quad (16)$$

Since $\frac{\partial^2(EC)}{\partial Q^2} > 0$, solving $\frac{\partial(EC)}{\partial Q} = 0$ gives Lemma 1.

The time complexity of inducing π_f^* is $O(1)$, while the time complexity of determining π^* is $O(M^2)$. According to Lemma 1, all the decisions in π_f^* can be pre-determined before the start of the time horizon using (14), while the decisions in π^* depend on current system states. Hence, the implementation of π_f^* is more computational efficient than the implementation of π^* .

Lemma 2. $EC_{\pi_f^*} = EC_{\pi^*}$, if the following condition is satisfied:

$$\frac{C_k \sum_{j=1}^n \lambda_{q_j}}{l \min(c_q) \lambda_d} < 1 \quad (17)$$

where $\min(c_q)$ is the minimum cost among c_{q_i} , for $i = 1, 2, \dots, n$.

Proof. Under condition (17), we get, $C_k < \frac{\lambda_d}{\sum_{j=1}^n \lambda_{q_j}} l \min(c_q)$ (18)

$\min(l_m)$, the minimum amount of knowledge loss at decision point m , for $1 \leq m \leq M$, is

$$\min(l_m) = \frac{\lambda_d}{\sum_{j=1}^n \lambda_{q_j}} l \quad (19)$$

Multiplying (19) with $\min(c_q)$, we get $\min(c_m^l)$, the minimum cost of knowledge loss at decision

point m , for $1 \leq m \leq M$, $\min(c_m^l) = \frac{\lambda_d}{\sum_{j=1}^n \lambda_{q_j}} l \min(c_q)$ (20)

According to (20) and (18), we have $C_k < \min(c_m^l)$, for $1 \leq m \leq M$, which means that, at any decision point over the time horizon, the cost of running KDD is less than the minimum cost of knowledge loss at the decision point. Hence, under condition (17), the optimal knowledge refreshing policy, π^* , requires KDD to be run whenever there is a query to the knowledge base, which is also a fixed knowledge refreshing policy. Therefore, $EC_{\pi_f^*} = EC_{\pi^*}$, if condition (1) is satisfied.

FUTURE RESEARCH

Future works need to be done in the following areas: (1) the magnitude of cost saving between EC_{π^*} and $EC_{\pi_f^*}$ needs to be studied both analytically and numerically; (2) the method to integrate when to refresh knowledge (i.e. this research) with how to refresh knowledge (i.e., data streaming research) needs to be developed.

REFERENCES

- [1] Can, F. 1993. Incremental clustering for dynamic information processing. *ACM Transactions on Information Systems*. 11(2) 143-164.
- [2] Fayyad, U., G. Piatetsky-Shapiro, P. Smyth. 1996. The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*. 39(11) 27-34.
- [3] Ganti, V., J. Gehrke, R. Ramakrishnan, W. Y. Loh. 1999. A framework for measuring changes in data characteristics. *Proc. of the 18th ACM SIGMOD Symposium on Principles of Database Systems*, 126-137.
- [4] Guha, S., A. Meyerson, N. Mishra, R. Motwani, L. O'Callaghan. 2003. Clustering data streams: theory and practice. *IEEE Transactions on Knowledge and Data Engineering* 15(3): 515-528.
- [5] Manku, G. S., R. Motwani. 2002. Approximate frequency counts over data streams. *Proc. of the 28th VLDB Conference*.

Selecting decision tree algorithms in the presence of noise

Authors:

Roger Blake (corresponding)
Lecturer
University of Massachusetts, Boston
Boston, MA 02125
617-287-7692
roger.blake@umb.edu

Paul Mangiameli
Professor
University of Rhode Island
Kingston, RI 02881
(401) 874-4217
mangia@uri.edu

Selecting decision tree algorithms in the presence of noise

Abstract

This paper investigates how selecting a classification or regression tree can lead to higher levels of predictive accuracy when an attribute variable has random or systematic noise. Such a selection is viable in a large number of applications, and this paper analyzes how to make the best selection when either type of attribute noise is present. Continuous class variables are often predicted using regression trees. An alternative is to discretize the continuous class variable and solve the problem using a classification tree. This research compares these two approaches when an attribute variable has either random or systematic noise. Few studies have considered the impact of attribute noise on decision tree selection, and none has analyzed how regression and classification performance differ when either random or attribute noise is present. The results show that especially in cases of systematic attribute noise, a classification tree approach can outperform a regression tree and is the best choice. Except for minor cases, the results show that for random attribute noise a regression tree is the more accurate and best choice. A model demonstrates how to make the best selection of a classification or regression tree in practice.

1. Introduction

This research evaluates how random and systematic attribute noise affect the performance of regression and classification trees, and analyzes how to use these differences for selecting the decision tree algorithm for highest performance.

Regression and classification trees have not been compared often; a regression tree is typically used for continuous class variables, and classification tree variables. To make this comparison, a continuous class variable needs discretization for the classification algorithm, losing a degree of information and reducing the appeal of this comparison. Yet many real-world problems exist in which this comparison is feasible. We believe that when those problems have conditions of random or systematic attribute noise, there are performance differences between regression and classification that make this comparison not only feasible, but which can be used to select the type of decision tree algorithm which leads to better performance.

The impact of noise has been an important topic for years, and well studied for decision tree algorithms. Studies of the impact of noise on decision trees have placed much more emphasis on noise in class, rather than attribute, variables. These studies have typically considered noise as random, and for continuous variables as randomly distributed perturbations. While random noise is more conducive to analysis, real-world data often includes systematic noise, which has rarely received direct attention.

This research considers attribute noise, and attribute noise that is either random or systematic. It evaluates how this noise affects the comparative performance of regression and classification trees. This is a new approach, and it is used to provide a method by which selecting between regression and classification in conditions of attribute noise will lead to performance improvements.

Presented next is prior work in this area, followed by our methodology, results, a method for decision tree selection, and a summary.

2. Prior work

The first section of prior work describes problems in which a comparison of classification and regression was viable; these occur in many fields, and are not always solved with decision trees. Prior research on the method used to make this comparison, discretization, follows. The last two sections describe prior work on noise, and then for decision tree noise.

2.1 Problem areas for comparing classification and regression

Problems often involve predicting a continuous variable. There are many real-world problems in which the usefulness of those predictions are the same whether expressed as a single figure, or expressed as falling into a range of figures. Many examples occur

when predictions of a continuous variable are used in a decision-making process. Decisions often use predicted values in combination with other factors such as cost, benefit, constraints, and qualitative considerations.

When predictions are equally useful when expressed as point values or as ranges, then it is feasible to discretize the class or dependent variable and use classification. In doing so, a comparison of regression and classification performance can be made. The real-world problems in which this approach is feasible are numerous. For example, Bodapati and Gupta (2004) describe the need to translate continuous predictions into categorical predictions as "a problem that occurs frequently" in marketing.

2.2 Discretization of class variables

In order to compare regression and classification trees, a continuous class variable must be discretized so that the classification algorithm can be applied. This is in contrast to a more frequent type of discretization, that of an attribute variable, often done to reduce computational burden.

Discretization methods are categorized as either supervised or unsupervised (Liu 2002); some supervised methods are statistically-based, such as the Chimerge algorithm developed by Kerber (1992) and its many variations. Supervised discretization methods in general use criteria such as reducing entropy or maximizing information gain as in the algorithms proposed by Fayyad and Irani (1992), or described by Dougherty, Kohavi and Sahami (1995).

However, if it is a class variable to be discretized, a supervised discretization technique is not feasible. Unsupervised discretization, which is accomplished by binning (Liu 2002), is the only option when class values are not known or do not exist (Witten and Frank 2000). Binning uses a pre-determined number of bins, and proceeds by grouping continuous data into either equal-width or equal-frequency bins (Liu 2002). Some equal-width binning methods simply divide the maximum and minimum of a continuous variable into even intervals. Other binning methods, such as those using the 3-4-5 rule, develop equal-width bins with the aim of intuitive interpretation (Han and Kamber 2000).

Equal-width binning is the discretization method that can be made to guarantee the retention of semantic meaning, because the number of equal-width intervals is taken as a parameter and the number of bins can be made small enough to satisfy this constraint. Therefore, equal-width binning was chosen as the discretization technique for this evaluation.

There have been few studies of discretizing a class variable to enable comparison with a classification algorithm. Bodapati and Gupta (2004) did so for direct marketing response data, finding some success, although they did not utilize decision trees. Peng and Flach (2001) found discretizing class variables a useful step to build fuzzy decision trees although did not consider noise. Because discretization loses a degree of

information, in general this approach has not been considered beneficial for class variables (Quinlan 1992).

2.3 Noise

Noise is defined as anything which obscures the relationship between attributes and class (Hickey 1996), and is considered to be present whenever incorrect class or attribute data is inconsistent with benchmark theories (Zhu and Wu 2004). Noise can originate from random sources, in which noise in continuous data takes the form of normally distributed perturbations, or noise can arise systematically (Elnahrawy and Nath (2003), taking the form of skewed perturbations, for which a normal distribution may not be a reasonable approximation.

For real-world applications, whether random systematic, the impact of noise is vitally important. Noise levels above 5% have often been found in applications (Zhu and Wu 2004), and noise levels of 10% have been considered unavoidable (Aze and Kodratoff 2002). In some cases, noise has been found to exceed 20% (Fayyad et al. 1996).

2.4 Decision trees and noise

Many approaches to noise developed for decision tree analysis have attempted to identify and remove it at different steps of decision tree analysis. For example, Gamberger, Lavrac and Dzeroski (1996) developed a method of instance elimination prior to tree induction by creating sets of training data for which a complete hypothesis could be constructed, i.e. had no contradictory cases, and then discarding instances by using a compression measure. John (1995) developed ROBUST-C4.5, which uses C4.5 to iteratively reclassify datasets after all misclassifications are identified and removed. In a sense, ROBUST-C4.5 is a wrapper to C4.5 and removes misclassifications prior to each iteration and re-classification by C4.5.

However, there is no guarantee that identifying and eliminating noise will improve decision tree performance. Oates and Jensen (1997) empirically determined that methods which eliminate instances provide no performance improvements. They analyzed algorithms with different tree pruning methods, specifically including those that are error-based (the C4.5 default), reduced error, minimum descriptive length, and cost complexity methods. The size of tree and performance from each algorithm was evaluated as the number of instances in the training datasets were randomly removed from full datasets, or added to an empty dataset. Each algorithm studied showed a rapid increase in performance as more instances were added randomly until there were approximately 200 instances (or about 15% of the total number of instances), and leveled off thereafter. This analysis used their threshold figure to generate synthetic data, discussed later. Oates and Jensen concluded that removing instances, either randomly or by filtering, might be beneficial in reducing tree size, but not in improving performance.

Moreover, efforts to develop more noise tolerant algorithms have focused on class variables rather than attribute variables (Barutcuoglu and Alpaydin 2003; 2003; Zhu and Wu 2004). One reason may be that class noise has been more readily acknowledged to have an adverse effect on performance (Quinlan 1986; Dietterich 2000; Dietterich 2002), so there has been more incentive to study class rather than attribute noise. Zhu and Wu (2004) cite this as one reason why, with the exception of missing values, “research on handling attribute noise has not made much progress.” But attribute noise is an important consideration. Zhu and Wu attribute noise more prevalent than class noise in practice. As they stated, “An interesting fact from real-world data is that the class information is usually much cleaner than what we thought; and it is the attributes that usually need to be cleaned.”

Among relatively few studies of attribute noise, Ling and Yan (2003) injected randomly generated noise into both class and attribute variables as part of evaluating a new approach which pruned trees by calculating probability estimates at every leaf node. They found improvements in their approach when compared to C4.5, however, they did not explicitly consider the impact of varying levels of random noise, systematic noise, or discretizing the class variable for classification.

In a study of attribute noise made Kalapanidas et al. (2003) injected random noise into both synthetic and real-world datasets. They varied attribute noise levels from 0% to 50%, and evaluated discretized a class variable for comparison with classification algorithms. The authors discretized the class variable into 5 bins, and found that in conditions of very high attribute noise, a classification tree could sometimes outperform a regression tree. This is consistent with some of the results of this study. However, they did not vary the number of bins used for discretization, nor did they consider systematic noise. They also did not develop any methods to choose either type of tree.

The majority of research on the impact of noise on decision trees has considered class noise instead of attribute noise. Research has almost invariably studied random noise and not systematic noise. Yet noise can appear in attributes as well as class variables, and noise can be either random or systematic. Although many potential applications exist, researchers have considered little benefit in discretizing a class variable for the purposes of comparing classification with regression. In situations without noise, that may be true, but the introduction of noise, particularly systematic noise, might change this.

3. Methodology

3.1 Hypotheses

To be able to determine how to select a classification or regression for best performance when noise is present, we first need to ascertain if there are significant performance differences between the two. We do this comparison for cases in which the number of bins used for discretization is five or fewer, in order to represent situations most likely to occur in reality. To determine if such differences exist, we state our formal hypotheses for random and systematic attribute noise separately, each including two corollary hypotheses to its alternate:

Hypothesis I: When varying levels of random (normally distributed) attribute noise are present, there will be no significant difference in the performance of regression and classification trees, when this comparison is based on five or less bins for discretization.

Corollary Hypothesis IA: If significant differences in algorithm performance do exist between regression and classification trees when random noise is present, they will vary depending on both the variability of the class variable and the level of attribute noise, when this comparison is based on five or less bins for discretization.

Corollary Hypothesis IB: If significant differences in algorithm performance do exist between regression and classification trees when random noise is present, they will vary depending on both the number of bins used for discretizing the class variable and the level of random noise, when the number of bins is five or less.

Hypotheses IA and IB are corollaries to the alternate of hypothesis I, and is to be evaluated only if hypothesis I is rejected. This corollary hypothesis considers the effects of class variability and random attribute noise on differences in performance. Hypothesis IB considers the effects of the number of bins chosen for discretization and the level of random attribute noise on performance differences. For regression and classification trees considered individually, and using the midpoint value within each bin as a class variable for a regression tree, the number of bins will affect performance. In the extreme case in which 1 bin is used, all class values will be discretized into the same bin, resulting in 100% correct predictions regardless of attribute noise. In opposite cases in which an increasingly large number of bins are used, accuracy will decrease and ultimately approach 0%, regardless of noise. Hypothesis IB is concerned with whether or not random attribute noise levels and the number of bins have significant effects on the differences between regression and classification performance.

The hypotheses for systematic noise are analogous to those for random noise:

Hypothesis II: When varying levels of systematic (skewed) attribute noise are present, there will be no significant difference in the performance of regression and classification trees, when this comparison is based five or less bins for discretization.

Corollary Hypothesis IIA: If significant differences in algorithm performance do exist between regression and classification trees when systematic noise is present, they will vary depending on both the variance of the class variable and the level of attribute noise, when this comparison is based on five or less bins for discretization.

Corollary Hypothesis IIB: If significant differences in algorithm performance do exist between regression and classification trees when systematic noise is present, they will vary depending on both the number of bins used for discretizing the class variable and the level of attribute noise, when the number of bins is five or less.

To compare classification and regression tree performance, a consistent metric is necessary. For classification trees, a frequently used measure of performance is the proportion of misclassified instances (Turney 1995), and is appropriate here.

3.2 Experimental design

Synthetic datasets with random attribute noise were generated to evaluate hypothesis I and its corollaries other synthetic datasets with systematic attribute noise were generated to evaluate hypothesis II and its corollaries.

Datasets consisted of 500 instances of a continuous class variable and a continuous attribute variable with the specifics provided below. The particular size of 500 instances for each dataset was chosen because Oates and Jensen (1997) concluded that somewhat over 200 instances was near a performance threshold for decision trees. Their work showed no performance improvements when dataset sizes were increased substantially beyond these levels.

A regression and then classification algorithm was used to analyze each dataset, and the difference in accuracy between them was calculated. This difference in performance for a dataset, expressed as a percentage, was a single data point for the comparison analysis in this study. To evaluate performance differences for random noise, datasets were created for distinct combinations of class variability, attribute noise level, and the number of bins used for discretization. For each these combinations, 300 datasets were generated.

Continuous class variables were generated using a normal distribution to represent actual data. Datasets were generating for different levels of class variability, class variability defined as the coefficient of variation, or the standard deviation divided by

the mean and expressed as a fraction. Class variability was ranged from .10 through .50 in increments of .10, therefore producing 5 levels of class variability for analysis. This specific set of values simulates real-world data and is the same range as used by Kalapanidas et al. (2003).

Datasets with systematic attribute noise were generated for CV_A varying from .10 to .50 in increments of .10, and the number of bins ranging from 2 to 5, as was done for random noise. Each combination of class variability and bins was held constant for 49 combinations of α_A and β_A , each varying from .1 through 2.5 in increments of .4. This range of Beta parameters created systematic attribute noise with highly variable skewness, and with more diverse distributions for systematic noise than random noise, more generated datasets were required.

Systematic noise was created in combinations of 5 levels of class variability, 49 levels of systematic attribute noise, and 4 levels of bins. The 300 datasets created for each combination were analyzed to find the difference in performance between regression and classification; each performance difference was used as a single data point for comparison. All performance comparisons were made with the M5' regression and the J48 classification algorithms, as implemented in Weka (Witten and Frank 2005).

4. Results

4.1 Results for random noise

Hypothesis I states there will be no significant difference between the performance of regression and classification trees when random attribute noise is present. This was evaluated by conducting an ANOVA having 9 treatment levels of random attribute noise. Performance differences were expressed as percentages, with 1.00 equaling 1%, and used as the dependent variable. These differences were calculated by subtracting the classification tree performance from the regression performance; positive values represent better regression tree performance, and negative values better classification tree performance.

To find if these differences were significant, an ANOVA was conducted to evaluate the effects of random attribute noise, with results shown in Table 1.

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4978.165	8	622.271	199.093	.000
Within Groups	68750.380	53991	3.126		
Total	173728.54	53999			

Table 1: ANOVA to evaluate hypothesis I

It is expected that increasing noise levels will degrade the performance of any decision tree algorithm. However, the difference in performance between regression and classification algorithms is the important factor here. This ANOVA shows that random noise has a significant effect on performance differences, $F_{(8,53991)} = 199$, $p < .01$. We therefore reject hypothesis I.

Therefore, Hypothesis IA was evaluated by a series of two way ANOVAs with factors of class variance and random noise level, in which the number of bins was held constant. The results of each ANOVA showed significant main effects for attribute noise, in which all $F_{(8,13445)}$ values were greater than 40, $p < .01$, and also for class variance in which all $F_{(4,13445)}$ values were greater than 26, $p < .01$.

Figure 1 shows representative results of displaying mean performance differences as they vary by the level of random attribute noise. This plot, based on 4 bins for discretization, displays a line for each level of class variability.

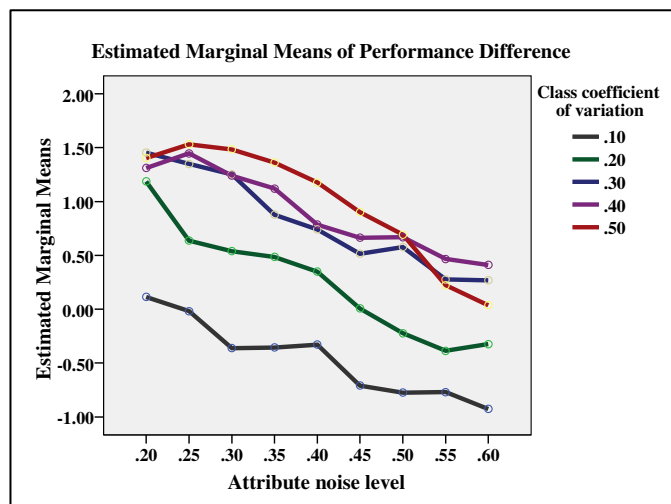


Figure 1: Means plots for each level of class variance held at 4 bins

Each line in Figure 1 represents mean performance differences for a single value of class variability, and shows how performance changes for random attribute noise. For example, the top line is shows how performance differences vary with attribute noise when class variability is held constant at .50, the bottom line how performance differences vary with attribute noise when class variability is held constant at .10. The mean performance difference for each highlighted point represents 300 datasets. This plot illustrates the significant differences in performance found in evaluating Hypothesis IA.

Hypothesis IB evaluated the effects of the number of bins used for discretization by a series of two-way ANOVAs in which class variability was held constant. Each of these ANOVAs showed significant effects of bins on differences in performance. A plot of representative data for these analyses is shown in Figure 2.

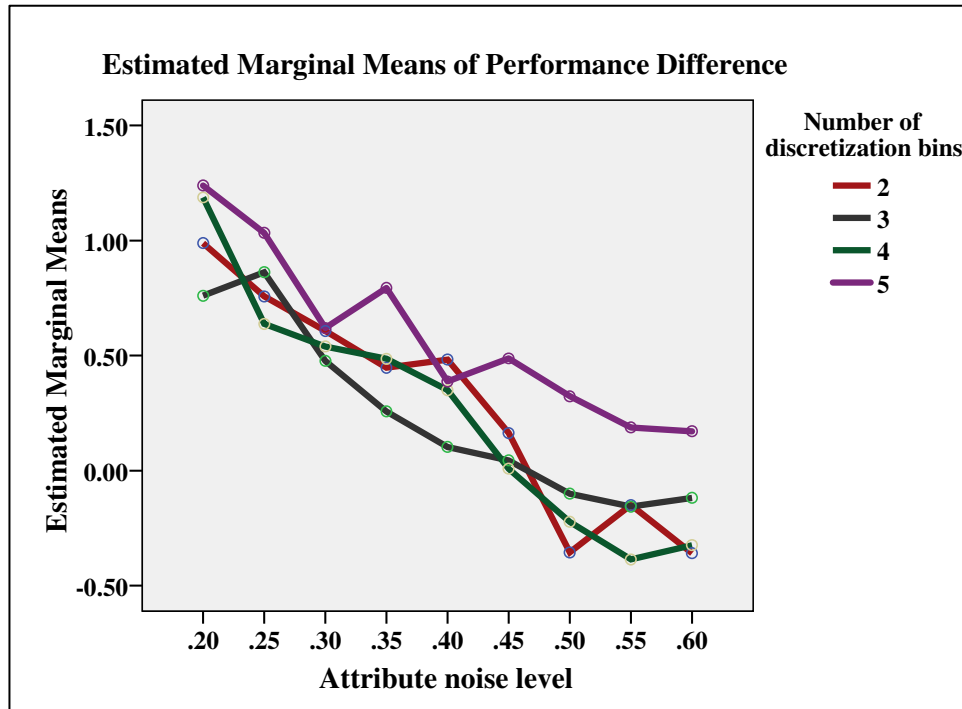


Figure 2: Means plots for each number of bins

Plots for other levels of random noise, from .10 to .60, had similar characteristics as Figure 2. With random noise, regression tends to perform better. Only when attribute noise reaches a high level, approximately .50, will classification be the better choice.

4.2 Results for systematic noise

In contrast to hypothesis I for random noise, hypothesis II focuses on determining if systematic noise has significant impacts on performance differences. Datasets with systematic noise were generated with varying combinations of Beta distribution α and β parameters. However, in practice it is improbable that the exact values of the α and β parameters of systematic noise are known in advance of selecting the best type of decision tree model, whereas having a sense of the skewness is more likely. Therefore, in order to improve the ability to apply these results, skewness was taken as representative of the level of systematic noise. Table 2 shows the ANOVA that evaluates the effects of systematic noise, skewness, on mean performance differences. This effect is significant; $F_{(50,293949)} = 5388$ and $p < .01$.

Tests of Between-Subjects Effects					
Dependent Variable: Performance difference between regression and classification					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	4856558.240 ^a	50	97131.165	5388.073	.000
Intercept	2955631.749	1	2955631.749	163955.197	.000
Skewness	4856558.240	50	97131.165	5388.073	.000
Error	5299039.076	293949	18.027		
Total	14786385.220	294000			
Corrected Total	10155597.315	293999			

^a. R Squared = .478 (Adjusted R Squared = .478)

Table 2: ANOVA results for effects of systematic noise on mean performance difference

Based on these results we reject hypothesis II and conclude that systematic noise has a significant effect on the mean performance difference between regression and classification trees. Accordingly, we proceeded to evaluate corollary hypotheses IIA and IIB for class variability and number of bins, repeated in the same fashion as the corollary to the alternate of hypotheses I.

The ANOVAs to evaluate hypothesis IIB were each also significant for each of the main effects, which were skewness and number of bins for discretization. All $F_{(50,58628)}$ values for the main effect of skewness were greater than 1,429 and $p < .01$, and all $F_{(3,58628)}$ values for the main effect of number of bins were greater than 1,318, and $p < .01$.

Although systematic attribute noise, class variability, and the number of bins all have significant effects on the performance difference of regression and classification, the nature of the relationships between each of them are much more complex than is the case for random noise. The general conclusion that can be made from examining the data and ANOVA results is that classification will have the higher performance when attribute noise is highly skewed. An example of this is shown in Figure 5, which plots the relationship between mean performance difference and skewness while holding the number of bins and class variability constant. Each point in the plot represents 300 observations.

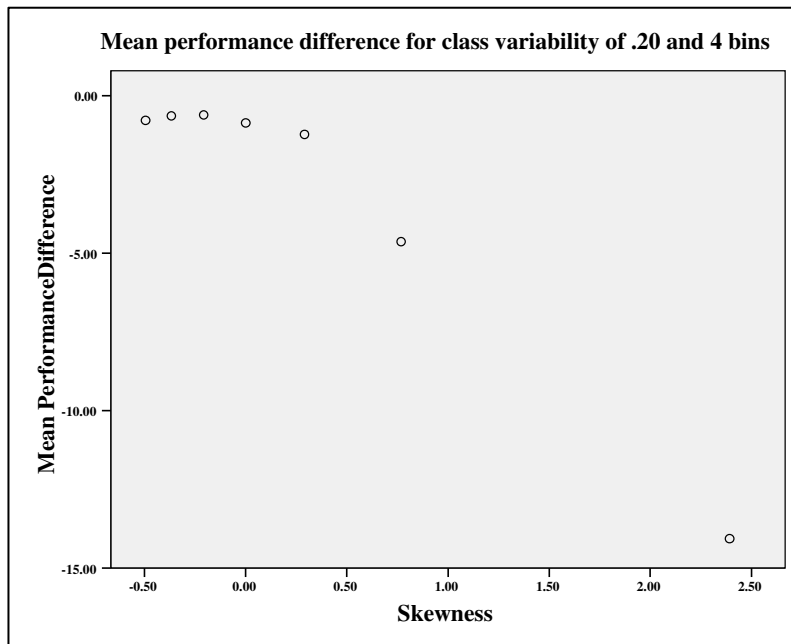


Figure 3: Mean performance differences holding bins constant at 4 and class variability at .20

Mean differences are closest to 0, indicating regression and classification are closer in performance, when skewness is near 0. As skewness increases, the tendency for classification to become the better choice increases, as seen by negative performance differences.

Regardless of whether these relationships are straightforward or complex, the objective of this research is to find quantifiable differences in performance under conditions of noise, and to enable the selection of a decision tree algorithm to take advantage of those differences. A method for this selection is the topic of the next section.

5. Choosing the right decision tree algorithm

Choosing the right decision tree under conditions of noise requires quantifying the results found in this study. If a mathematical approach were taken, the irregularity of systematic noise would introduce a high level of complexity, requiring a set of strong and restrictive assumptions. In such cases mathematical methods quickly become intractable (Piater et al. 1998). An alternative to a mathematical approach is to create a model of algorithm performance. Such a model was developed by Courtney, Thacker and Clark (1997) to analyze the effectiveness of computer vision methods for datasets injected with varying levels of noise, and the same general type of model as is constructed here.

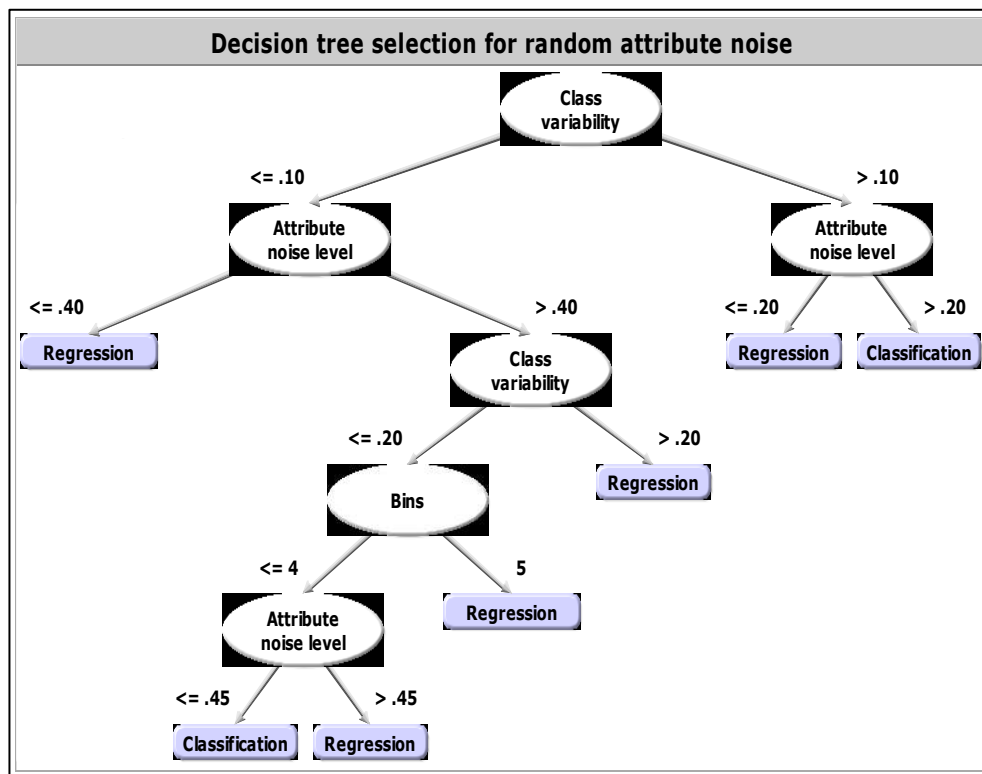
Separate models of algorithm performance for random and systematic attribute noise were built, since in practice the nature of attribute noise is likely to be known at least to this level. These models were decision trees created with the J48 classification tree algorithm. For each decision trees, the algorithm with the highest level of accuracy – either regression or classification - was used as the categorical class variable to be predicted, and the specific parameters used in synthesizing datasets (class variability, attribute noise level, number of bins) as the attributes.

In cases of random attribute noise, the best performing algorithm was correctly predicted in 95% of the test cases; overall precision and recall were both above 90% as shown in Table 3.

True Positive Rate	False Positive Rate	Precision	Recall	Class
0.964	0.095	0.971	0.964	Regression
0.905	0.036	0.884	0.905	Classification

Table 3: Confusion matrix for the model of algorithm performance for random noise

This decision tree for random attribute noise was reasonably compact, having 13 nodes of which 7 are leaves. This is so because there are only a few exceptions to regression trees being the better choice for random noise. Figure 4 shows the decision tree to select either regression or classification when random attribute noise is present.



The model of algorithm performance for systematic attribute noise was not as accurate as that for random attribute noise, but did predict whether regression or classification would be best in 91% of the test cases. As can be seen in the confusion matrix of Table 4, precision and recall rates were somewhat lower than for random attribute noise, but overall were and respectively.

True	False			
Positive Rate	Positive Rate	Precision	Recall	Class
0.962	0.458	0.935	0.962	Classification
0.542	0.038	0.674	0.542	Regression

Table 4: Confusion matrix for the model of algorithm performance for systematic noise

The decision tree to model algorithm selection with systematic noise was somewhat larger than that for random attribute noise, with 47 instead of 13 nodes and so is not shown here, but is equally useful.

The models of algorithm performance, developed as decision trees, have a high degree of predictive accuracy for both random and systematic attribute noise. With access to raw data, an analyst is routinely aware of the general type of attribute noise. Armed with this information and some of the basic characteristics of attribute noise, the variability of the class variable, and the number of bins appropriate for the application, an analyst can select the best decision tree to use.

6. Conclusions

6.1 Summary

This research demonstrates that there are significant differences in the performance of classification and regression trees when attribute noise is present. More importantly that those differences can be quantified and used to select the decision tree algorithm, either classification or regression, for highest performance. The performance differences found from analysis varied over a wide range, but even small differences in predictive accuracy can have significant importance in practice. This is true for direct marketing, insurance, and investments, and for a wide range of other applications including natural resource allocation, software fault detection, and medical diagnosis among many others.

These performance differences are reachable by using the algorithm models from the paper. By knowing only whether attribute noise random or systematic, and having basic characteristics of the attribute and class variables, and an understanding of the application, an analyst can use the algorithm models presented here. If classification is

indicated for better performance, the analyst needs to discretize the class variable; if regression is indicated, no additional pre-processing is necessary.

6.2 Limitations

There are several limitations to this research. As we have used synthetic data in our analysis, we are not yet able to generalize these findings to actual datasets. Datasets in the real-world in some domains may have different, and possibly more systematic, noise than we have created in our datasets, and should be investigated.

Furthermore, each of the datasets we generated consisted of only one attribute variable. We did not investigate an auto-correlated attribute, as may exist in some data. While we therefore isolated any possible interactions between attributes, realistic datasets may have such interactions. But the impact of attribute noise is related to the attribute's correlation with the class variable (Zhu and Wu 2004). We did not investigate auto-correlated attributes, as from a time series. Finally, we only considered continuous and not categorical attributes. Further research towards evaluating each of these would extend applications of selecting classification or regression for best performance.

6.3 Further research

Our analysis used the J48 algorithm for classification and the M5' algorithm for regression. There are many variations of bagging, boosting, and other ensemble methods available, and a great many other classification and regression algorithms available as well. More research could investigate how the finding of this paper may, or may not, with ensemble techniques or different decision tree algorithms. Other research could determine if other machine learning techniques such as neural networks would yield similar results.

These areas for further research are promising; each might identify more opportunities in which performance gains can be identified, quantified, and put into practice through the selection of a specific algorithm in response to varying conditions of noisy data.

Bibliography

- Aze, J. and Y. Kodratoff (2002). "A Study of the Effect of Noisy Data in Rule Extraction Systems." Proceedings of EMCSR 2: 781-786.
- Barutcuoglu, Z. and E. Alpaydin (2003). A Comparison of Model Aggregation Methods for Regression. Artificial neural networks and neural information processing - ICANN / ICONIP 2003. Istanbul, Turkey. **2714**: 76-83.
- Bodapati, A. and S. Gupta (2004). "A Direct Approach to Predicting Discretized Response in Target Marketing." JOURNAL OF MARKETING RESEARCH.
- Courtney, P., N. A. Thacker and A. F. Clark (1997). "Algorithmic modelling for performance evaluation." Machine Vision and Applications 9: 219-228.
- Dietterich, T. (2000). "Ensemble Methods in Machine Learning." Lecture Notes in Computer Science 1857: 1-15.
- Dietterich, T. G. (2002). "Ensemble learning." The Handbook of Brain Theory and Neural Networks 2.
- Dougherty, J., R. Kohavi and M. Sahami (1995). Supervised and Unsupervised Discretization of Continuous Features. International Conference on Machine Learning.
- Elnahrawy, E. and B. Nath (2003). "Cleaning and querying noisy sensors." Proceedings of the 2nd ACM international conference on Wireless sensor networks and applications: 78-87.
- Fayyad, U. M. and K. B. Irani (1992). "On the Handling of Continuous-Valued Attributes in Decision Tree Generation." Machine Learning V8(1): 87-102.
- Fayyad, U. M., G. Piatetsky-Shapiro and P. Smyth (1996). "Knowledge Discovery and Data Mining: Towards a Unifying Framework." Knowledge Discovery and Data Mining: 82-88.
- Gamberger, D., N. Lavrac and S. Dzeroski (1996). "Noise elimination in inductive concept learning: A case study in medical diagnosis." Proc. of the 7th International Workshop on Algorithmic Learning Theory: 199-212.
- Han, J. and M. Kamber (2000). Data Mining: Concepts and Techniques, Morgan Kaufmann.
- Hickey, R. (1996). "Noise modelling and evaluating learning from examples." Artificial Intelligence 82(1-2): 157-179.

- John, G. H. (1995). "Robust decision trees: Removing outliers from databases." Proceedings of the First International Conference on Knowledge Discovery and Data Mining: 174–179.
- Kalapanidas, E., N. Avouris, M. Craciun and D. Neagu (2003). Machine Learning algorithms: a study on noise sensitivity. Proc. 1st Balcan Conference in Informatics Thessaloniki.
- Kerber, R. (1992). "ChiMerge: Discretization of Numeric Attributes." National Conference on Artificial Intelligence: 123-128.
- Ling, C. X. and R. J. Yan (2003). "Decision tree with better ranking." Proceedings of the 20th International Conference on Machine Learning. Morgan Kaufmann: 480-487.
- Liu, H., Hussain, F., Tan, C., Dash, M. (2002). "Discretization: An enabling technique." Journal of Data Mining and Knowledge Discovery **6**: 393-423.
- Oates, T. and D. Jensen (1997). "The effects of training set size on decision tree complexity." Proceedings of the Fourteenth International Conference on Machine Learning **254261**.
- Peng, Y. and P. A. Flach (2001). "Soft Discretization to Enhance the Continuous Decision Tree Induction." Integrating Aspects of Data Mining, Decision Support and Meta-Learning (PKDD'01 Workshop): 109–118.
- Piater, J. H., P. R. Cohen, X. Zhang and M. Atighetchi (1998). "A randomized ANOVA procedure for comparing performance curves." Proc. 15th International Conf. on Machine Learning: 430-438.
- Quinlan, J. R. (1986). "Induction of decision trees." Machine Learning **1**(1): 81-106.
- Quinlan, J. R. (1992). Learning with Continuous Classes. 5th Australian Joint Conference on Artificial Intelligence.
- Turney, P. D. (1995). "Cost-Sensitive Classification: Empirical Evaluation of a Hybrid Genetic Decision Tree Induction Algorithm." Journal of Artificial Intelligence Research **2**: 369-409.
- Witten, I. and E. Frank (2000). Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations, Morgan Kaufman.
- Witten, I. and E. Frank (2005). Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations. San Francisco, Morgan Kaufman.

Zhu, X. and X. Wu (2004). "Class Noise vs. Attribute Noise: A Quantitative Study."
Artificial Intelligence Review **22**(3): 177-210.

MODELING INDIVIDUAL CONSUMER PREFERENCES USING CONJOINT ANALYSIS AND NEURAL NETWORKS

Frank LoPinto
Management Department, La Salle University
1900 West Olney Ave., Philadelphia, PA 19141
Phone: (215) 951-1517, E-mail: lopinto@lasalle.edu

Cliff T. Ragsdale
Department of Business Information Technology
Virginia Polytechnic Institute and State University
1007 Pamplin Hall, Blacksburg, VA 24061
Phone: (540) 231-4697, E-mail: cliff.ragsdale@vt.edu

ABSTRACT

There is a need for better product searching tools on the Web. Online product search could be greatly improved if it was possible to quickly and easily capture and model individual consumer preferences. Conjoint analysis is a widely accepted technique used to elicit preferences of aggregate consumer groups, however, traditional conjoint analysis techniques have not been widely adopted for predicting individual consumer preferences. This paper presents a methodology for modeling individual customer preferences using conjoint analysis and neural networks. In this paper, preferences are elicited, preference models of individual participants are constructed, and the prediction accuracy of the methodology is measured.

Keywords: Preference modeling, conjoint analysis, neural networks

INTRODUCTION

The ever expanding number of merchants and products available on the World Wide Web (Web) continues to increase the effort, time, and uncertainty involved with business-to-consumer (B2C) online purchasing. As B2C e-commerce matures, the need for better product searching and purchasing tools increases. Many Web sites currently implement user-driven sorting and filtering to aid consumers with the difficult task of finding the right product for their needs.

Many researchers look toward software agents to perform these searching, negotiation, and purchasing tasks in the future [1][2][4][8][9]. Much research has been pursued on using agents in a distributed environment, yet little has been done to realistically elicit and model the preferences of consumers so that an online agent can truly act on a consumer's behalf. Traditionally, agent developers acquire knowledge from domain experts, then program the agent to make autonomous and/or intelligent decisions. This process requires significant time and effort on the part of the user, developer, and domain experts, and is impractical to perform individually for every consumer in the fast-paced global environment of the Web. It is evident that a fast, effective, and easy method for eliciting and modeling individual consumer preferences is needed for future e-commerce markets.

In this study, we construct and test the accuracy of preference models for individual consumers. CA, a technique primarily used in marketing to elicit and model target market preferences, was used to elicit the preferences of potential customers. However, traditional CA alone did not provide suitable preference models to adequately predict individual customer preferences. Therefore, we developed a methodology for modeling customer preferences using neural networks, and demonstrated the superiority of this new approach versus a traditional CA method in the context of spring break vacation package purchasing.

Section 2 of this paper provides background on software agents, CA, and neural networks. Section 3 outlines our new methodology for constructing user preference models using neural networks. Section 4 presents our analysis and results, and in Section 5 we discuss our conclusions and future work.

BACKGROUND

Conjoint Analysis and Neural Networks

CA is the method most used by marketing researchers and practitioners for analyzing consumer group trade-offs and eliciting and modeling product attribute preferences [6]. CA presents individuals in target consumer groups different (but similar) product descriptions / profiles with various attributes and asks them to either rank or rate the items. CA then uses quantitative techniques to estimate the structure of the target group's preferences [7]. Desirable product attributes can be determined from this information.

Artificial Neural Networks, or neural networks (NNs), are biologically-inspired computer-based systems that are modeled after the brain and nervous system [3]. The strength and allure of neural networks is that they rely on pattern recognition for determining a response to input stimuli rather than domain specific logic and rules required by expert systems. To date, NNs have not been used successfully on individual consumers in conjunction with CA because NNs require large data sets for training purposes. Therefore, in order to apply neural networks to the construction of an individual's preference model, we must develop a methodology for increasing the number of data records available from the data collected via CA.

METHODOLOGY

Overview

In this study we collected consumer preference data using a CA technique, constructed several models for individual participants' preferences, and compared the models in terms of predictive accuracy. This section details the data collection and experimental methodology.

Data Collection

Participants in this study completed a CA instrument designed to assess individual preferences for spring break vacation packages. This topic was chosen for its familiarity and relevance to its participants - fall semester senior level undergraduate students. We collected our data using a

full-profile CA. Using the full-profile method, participants are shown “cards” that describe complete vacation packages and are asked to rate each card. A card might be represented as follows: An all inclusive vacation to Mexico, at a 3-star hotel, traveling by air, at a price of \$2,500.

Mathematically, a card is defined as follows where a_{k_i} represents the level of attribute a_k contained in card c_i :

$$c_i = (a_{1_i}, a_{2_i}, \dots, a_{n_i}) \quad (1)$$

Each participant was provided a paper survey containing the same unique 25 spring break vacation packages (cards) and asked to perform two tasks. First, each participant was asked to assign the most preferable package a rating of 100, and the least preferable a rating of 0. Next, they were instructed to assign a unique rating from 1 to 99 for each of the remaining 23 vacation packages. Therefore, for each card, c_i , each participant provided a unique individual rating r_i such that $r_i \neq r_j \forall i, j \in [1, 25]$.

It has been shown that the number of stimuli should be limited to no more than thirty cards for full-profile CA data collection [7][10]. As a result, we employed a 5 x 5 orthogonal fractional factorial design requiring that 25 stimuli be presented to each participant. Surveys were collected from ninety-nine participants. Of the original ninety-nine, seven improperly completed the study, therefore leaving ninety-two completed questionnaires. As an incentive to take the rating task seriously, participants were told that \$100 cash award would be given to the participant whose preferences could be modeled most accurately.

ANALYSIS AND RESULTS

Overview

Once collected, the data were used to build four different preference models for each participant; three using regression and three using a neural network. We refer to these four models as R-23, N-23, R-253, and N-253.

The letter portion of the model name designates whether regression (R) or a neural network (N) was used to construct the model. The number portion of the model name represents the number of records used to construct the model. The regression models were constructed using Microsoft Excel to calculate the least squares parameter estimates for each model. The NNs were constructed using Neuralworks Predict version 3.12 by NeuralWare. Default settings were used for NN model construction and selection of the training, validation, and testing sets.

Evaluation

For each participant, a v-fold cross validation procedure was used for each model type. This involved constructing a total of 300 models for each participant using a pairwise “hold out” procedure in order to assess the effectiveness of each modeling technique in an unbiased manner. For each model, two unique cards, c_i and c_j , where $i \neq j$, were removed from the original data set of 25 cards prior to model construction. The model was then constructed using the remaining 23

cards. Repeating this for every unique combination of i and j resulted in $\binom{25}{2} = 300$ unique combinations of hold out cards, and subsequently 300 models and tests.

Each model was then used to predict a relative preference ordering between hold out cards c_i and c_j . In general, a given participant prefers c_i over c_j if $r_i > r_j$, and prefers c_j over c_i if $r_j > r_i$. The model's prediction, \hat{r}_i and \hat{r}_j , was then compared to the participant's actual ratings of r_i and r_j to determine accuracy. Predicting relative preference in this way is consistent with the philosophy of ordinal optimization. Fu [5] states that, "It is generally easier to compare solutions and find relative ordering among them rather than it is to estimate them precisely." That is, by predicting \hat{r}_i and \hat{r}_j individually and then comparing values, a relative preference ordering between cards c_i and c_j can be determined and compared to the participant's specified preference indicated by r_i vs. r_j .

Models

R-23 vs. N-23: The R-23 and N-23 models were initially intended to provide a prediction, \hat{r}_i , of the actual precise rating value, r_i , for each "hold out" card for each participant. Preliminary results showed that the models performed poorly for predicting ratings, but did a fair job at ordering, or determining the relative preferences of pair-wise comparisons.

For the R-23 and N-23 models, each record used for model construction is in the form $[c_i, r_i]$ where i is the original data record card number ($1 \leq i \leq 25$), c_i represents the card as described in above, and r_i is the unique numeric rating provided by each participant.

Using the predicted numeric ratings \hat{r}_i and \hat{r}_j for cards c_i and c_j , the preference determination rules were as follows:

If $\hat{r}_i > \hat{r}_j$ and $r_i > r_j$, then the model accurately predicted the preference.

If $\hat{r}_i \leq \hat{r}_j$ and $r_i < r_j$, then the model accurately predicted the preference.

For each subject, we computed the percentage of trials in which each of the 300 models made an accurate prediction. Table 1 summarizes the overall prediction accuracy for these two models for all 92 subjects.

R-253 vs. N-253: Although the N-23 models outperformed the R-23 models above, recall, that only 23 records were used to train the N-23 model. Because NNs typically perform better when trained with larger data sets we developed a method for modeling preference orders that modifies the original data and increases the number of training records. In R-253 and N-253, the original data was expanded by creating a single record for every pairwise comparison in the model-building card set. As in R-23 and N-23, a v-fold cross validation procedure was used such that 300 different models were constructed per participant for every card pair c_i and c_j , where $i \neq j$ such that c_i and c_j were excluded from construction of the model.

Therefore, each record used in the model took the form $[c_i, c_j, y_{ij}]$ where i and j are the original data record card numbers ($1 \leq i \leq 24, 2 \leq j \leq 25$), and $i < j$ for every record, c_i and c_j represent the

respective cards, and y_{ij} is a binary value representing the participant's preference between c_i and c_j as follows:

If $r_i > r_j$, participant prefers c_i and $y_{ij} = 0$

If $r_i < r_j$, participant prefers c_j and $y_{ij} = 1$

Note that $r_i \neq r_j$, since participants were instructed to provide unique ratings for each card. The resulting models each contained 253 records. The model output, \hat{y}'_{ij} , took on the form of a real number from 0 to 1. The value of 0.5 was used as the cutoff such that:

If $\hat{y}'_{ij} < 0.5$, c_i is predicted as preferred over c_j , and $\hat{y}_{ij} = 0$

If $\hat{y}'_{ij} \geq 0.5$, c_j is predicted as preferred over c_i , and $\hat{y}_{ij} = 1$.

Once again, 300 unique models were constructed using the v-fold cross validation procedure. For each model, \hat{y}_{ij} and y_{ij} were calculated for the comparison between the held out cards c_i and c_j . Table 1 summarizes the prediction accuracy observed in this testing.

Again, three hundred unique models were constructed for this method using the same v-fold cross validation procedure described earlier. For each model, \hat{y}_{ij} , \hat{y}_{ji} , y_{ij} , and y_{ji} were calculated for the "held out" cards, c_i and c_j .

Table 1 provides a statistical summary of results for the entire study. Reported are the percent of times each model correctly predicted the preferences of each individual. In general, the neural network models outperform the regression models, and N-253 outperforms all other methodologies. Recall that R-23 represents the standard CA methodology. When comparing the results of the N-253 and R-23 techniques, average prediction accuracy is increased by 12.8% without a significant sacrifice in standard error. The most accurate model (N-253) provided correct predictions 97% (291 correct out of 300) of the time versus 88.33% for R-23, a 9.8% improvement. Additionally, the least accurate prediction increased from 21% with R-23 to 45% with N-253.

	R-23	R-253	N-23	N-253
Mean	0.6513	0.5703	0.7111	0.7344
Minimum	0.2100	0.3833	0.4200	0.4500
Maximum	0.8833	0.7100	0.9867	0.9700
Standard Error	0.0115	0.0071	0.0118	0.0119
Standard Deviation	0.1099	0.0678	0.1128	0.1137

Table 1: All results

CONCLUSIONS AND FUTURE WORK

Efficiently and effectively obtaining and modeling consumer preferences is a major roadblock in enabling agent-based negotiation on the Web. This paper suggests a promising approach to removing this roadblock and allowing the information superhighway to realize its full potential for B2C e-commerce. In this paper we examine the problem of modeling consumer preferences for agent-based B2C online purchase transactions. We introduce a class of problems that would benefit greatly from the addition of a system that could easily and accurately elicit and model individual customer preferences. Addressing solutions to these problems, we propose CA as the elicitation tool and conduct a study using 92 participants. While conducting a standard CA regression analysis of the data, we propose using a neural network as an alternative method for

modeling individual customer preferences and propose a number of ways to expand the original CA data set for use in a NN.

For both N-23 vs. R-23 and N253 vs. R-23, the mean, minimum, and maximum values (reported in Table 1) represent the proportion of times the models correctly predicted the participants' preference between cards c_i and c_j . For the 92 participants, the NN models again outperformed the regression models. A standard t-test for two-sample means indicated a significant difference in the means ($p \ll 0.0001$).

In comparing the proposed methodologies, the N-253 method proved to be the most accurate procedure. This methodology resulted in a significant improvement over its regression counterpart, R-253 while providing a more robust methodology than N-23 and N-253. In addition to outperforming a standard CA methodology our method also overcomes problems in currently proposed systems such as the requirement of domain experts to build a knowledge base, as well as the previously mentioned problem of sample size.

The positive results of our work present several opportunities to address additional questions as future work. One such question deals with the ability to instantiate an agent such that it will use the models developed here to act on behalf of a customer by either searching or negotiating for products.

REFERENCES

- [1] Berners-Lee, T., Hendler, J., & Lassila. "The semantic web." *Scientific American*, 2001, 284 (5), 34-43.
- [2] Bhargava, H. & Power, D. "Decision support systems and web technologies: a status report." *Seventh Americas Conference on Information Systems (AMCIS '01)*. Association for Information Systems. Boston, MA. 2001.
- [3] Burke, L. "Introduction to artificial neural systems for pattern recognition." *Computers and Operations Research*, 1991, 18(2), 211-220.
- [4] Fan, W., Gordon, M., & Pathak, P. "A generic ranking function discovery framework by genetic programming for information retrieval." *Information Processing & Management*. 2004, vol. 40(4), 587-602.
- [5] Fu, M. "Optimization for simulation: theory vs. practice." *INFORMS Journal on Computing*. 2002, 14(3), 192-215.
- [6] Green, P., Krieger, A., & Wind, Y. "Thirty years of conjoint analysis: reflections and prospects." *Interfaces*, 2001, 31(3), Part 2 of 2, S56-S73.
- [7] Green, P. & Srinivasan, V. "Conjoint analysis in consumer research: issues and outlook", *The Journal of Consumer Research*, 1978, 5(2), 103-123.
- [8] Guttman, R., Moukas, G., & Maes, P. "Agent-mediated electronic commerce: a survey." *The Knowledge Engineering Review*, 1998, 13(2), 143-152.
- [9] Lee W.-P., Liu C.-H., Lu C.-C. "Intelligent Agent-based Systems for Personalized Recommendations in Internet Commerce." *Expert Systems with Applications*, 2002, 22(4), 275-284.
- [10] Orme, B. Which conjoint method should I use? *Sawtooth Software Research Paper Series*. 2003

USING THE GAMMA MEMORY NEURAL NETWORK FOR BANKRUPTCY PREDICTION: A PRELIMINARY STUDY

Scott A. Dellana, East Carolina University, Greenville, NC 27858, (252) 328-4893, dellanas@ecu.edu
David West, East Carolina University, Greenville, NC 27858, (252) 328-6370, westd@ecu.edu

ABSTRACT

Many static neural networks have been studied extensively in financial classification problems. However, dynamic time series predictive classification using neural networks with memory, such as the Gamma Memory neural network (GMNN), may prove more accurate. In this study we compare the predictive accuracy of the GMNN to the Multilayer Perceptron neural network and the statistical approaches of Logistic Regression and Multiple Discriminant Analysis (MDA). Preliminary analysis of the GMNN versus other models gives mixed results.

INTRODUCTION AND LITERATURE REVIEW

The purpose of this research is to investigate if a time series classification method is able to predict potential bankruptcies more accurately than the static classification methods used since Altman's original Z-score model. We test a neural network time series classification model (Gamma memory neural network) and contrast the decision accuracy with several prominent static classification methods. To our knowledge, there is no published research on the application of time series models to bankruptcy prediction.

Bankruptcy prediction has been an area of interest for many decades. Altman (1968) studied bankruptcy prediction in his seminal work on MDA using financial ratios. Traditional methods of financial decision support include consumer credit scorecards (Brill, 1998; Henley, 1995; Mester, 1997; Reichert *et al.*, 1983; Rosenberg and Gleit, 1994) and discriminant models for assessing corporate financial health (Altman *et al.*, 1994; Reichert *et al.*, 1983). Both are essentially multivariate linear models that output a probability that the client will repay debt as agreed.

Neural networks have now come to the forefront as the preferred method for bankruptcy prediction (Kumar and Ravi, 2007). Many neural network architectures have been studied, including: the MLP neural network, the probabilistic neural network (PNN), the auto-associative neural network (AANN), self-organizing maps (SOM), learning vector quantization (LVQ) and cascade correlation (Cascor). Many of the studies involving these neural networks continue to compare them with statistical techniques such as factor analysis, Logit, and various forms of discriminant analysis. In many of the cases the neural networks are reported to provide more accurate bankruptcy prediction capability than the parametric statistical approaches. However, the results are also often mixed.

Tam and Kiang (1992), in an early treatment of neural networks in bankruptcy research, compared a variety of models. They studied MDA, Logit, K-nearest neighbor (KNN), a decision tree classification algorithm (ID3), a single-layer neural network, and a multi-layer neural

network. The neural networks used were standard back-propagation (BPNN). The multi-layer network was the best for predicting bankruptcy using financial ratios one year ahead of bankruptcy. For two years ahead of bankruptcy, Logit was the best in the same studies. Salchenberger, *et al.* (1992), when considering the bankruptcies of thrifts, found the BPNN significantly outperformed Logit. In a comparison of the BPNN with MDA, Coats and Fant (1993) found the BPNN to be generally better, although it had a wider variance in the classification result depending on the horizon used. Altman *et al.* (1994) considered 1000 Italian firms in a bankruptcy study that compared a BPNN and MDA. For a one-year-ahead prediction, MDA appeared to perform slightly better than the BPNN. Boritz and Kennedy (1995) compared a number of techniques, including different BPNN training procedures, Logit, and MDA. The results of the comparisons were inconclusive.

The BPNN has shown in many studies effectiveness in predicting bankruptcy that is mixed to good when compared to classical MDA and other approaches. Consequently, new hybrid techniques and genetic algorithms have recently begun to receive attention. Lee *et al.* (1996) tested combinations of the models such as MDA, ID3, self-organizing maps, and a BPNN. They studied bankruptcy prediction in Korean firms and reported that the self-organizing feature map assisted neural network performed best. Back *et al.* (1996) used genetic algorithms for input selection to a BPNN. The data in the study covered one to three years prior to bankruptcy. The BPNN was significantly better than MDA and Logit. Zhang *et al.* (1999) used a five-fold cross-validation scheme on a group manufacturing firms, comparing a BPNN and Logit for bankruptcy prediction. The BPNN significantly outperformed Logit. McKee and Greenstein (2000) developed an approach based on decision trees. They applied it to a sample of US firms for financial health data one year ahead of bankruptcy. Their method gave better results than MDA and BPNN for Type II error, but worse results for Type I error. Atiya (2001) developed novel indicators extracted from the equity markets for a neural network system. They showed that the use of these indicators, in addition to traditional financial ratio indicators, gave a significant improvement in the bankruptcy prediction accuracy for the neural network. Forecasts were based on financial information three years in advance of bankruptcy. Pendharkar (2005) studied a threshold-varying artificial neural network (TV-ANN) for solving the bankruptcy prediction problem. Using a set of simulated and real-world data sets, the TV-ANN performed well when compared to the BPNN and the MDA. The performance comparisons of TV-ANN with a genetic algorithm-based ANN and a classification tree approach gave mixed results. In a study by Lee, *et al.* (2005), supervised and unsupervised neural networks were compared using their representative types. The BPNN and the Kohonen self-organizing feature map, selected as the representative type for supervised and unsupervised neural networks, respectively, were compared for bankruptcy prediction accuracy. MDA and Logit were also performed to provide performance benchmarks. The findings suggest that the BPNN is a better choice when a target vector is available.

Lacking in all of these studies is an analysis of financial performance over time in making bankruptcy prediction. There is a reliance on information that is used to train the neural networks with data from one point in time – typically one or two years prior to bankruptcy. A time series approach would be desirable and could improve the accuracy of prediction. It has been shown in at least one application that the standard MLP, which is popular in bankruptcy prediction (Lee *et al.*, 2005), underperforms neural networks that are designed to manage temporal effects. One

such neural network is the GMNN. In a recent study of wastewater process control, which is characterized by nonlinear time varying dynamics, the Time Delay Neural Network -- a less sophisticated version of the more general GMNN -- was found to be superior to MLP in accuracy of prediction (Zhu *et al.*, 1998). There is, therefore, reason to suspect that the GMNN will provide a better prediction of bankruptcy than the MLP in a dynamic time series application.

DATA AND MODELS

Description of the Dataset

A small bankruptcy dataset has been constructed by the authors from Standard and Poor's Compustat financial files. Explanatory variables for this study include the five key financial ratios from research by Altman, *et al.* (1994). These ratios, constructed from financial statement information one to three years prior to bankruptcy include: working capital/total assets, retained earnings/total assets, earnings before interest and taxes/total assets, market value of equity/book value total liability, and sales/total assets. Data over a two-year period will be considered. The current dataset totals 45 companies, consisting of 23 bankrupt companies and 22 healthy companies. The financial information is quarterly data for each company. Ratios were obtained going back eight quarters starting one year prior to bankruptcy for the bankrupt companies. The same ratios were obtained for the healthy companies going back eight quarters from the most recent available data.

Multiple Discriminant Analysis

MDA can be used to classify an observation into one of several *a priori* groupings. This classification is based on the individual characteristics of the observation. MDA may be used to make predictions in problems where the dependent variable appears in qualitative form. MDA attempts to derive a linear combination of the characteristics which best discriminates among groups (Altman, 1968). In the case we consider, the groupings are bankrupt and non-bankrupt. MDA allows the analysis of the entire variable profile of the object simultaneously rather than sequentially examining its individual characteristics. In a financial health application, this allows combinations of financial ratios to be considered together (Altman, 1968). The reader is referred to Altman (1968) for a description of MDA in financial applications.

Logistic Regression (Logit)

Logit is a statistical model based on a discrete probability distribution which takes the value 1 with success probability p and value 0 with failure probability $q = 1 - p$. It is a generalized linear model that uses the Logit as its link function. Logit is essentially an MLP with no hidden layer, which is explained in the following section. The reader is referred to Agresti (2002) for a description of Logit model theory.

The MLP Neural Network

A neural network is a flexible, nonparametric modeling tool that builds a predictive model by learning to recognize patterns in historical data. Patterns in historical data can often exist in both spatial and temporal domains. The conventional MLP neural network, trained with the back-propagation of error algorithm, is designed to identify static patterns that do not vary over time. The MLP neural network is widely used in many applications, including financial health prediction. The basic structure of the feed-forward MLP neural network for financial

applications will consist of input nodes, the hidden layer, and two output nodes. External information (i.e., the financial ratios) is received at the input layer, while the output layer produces the model solution. In between is a hidden layer where patterns are identified. The three layer MLP neural network is commonly used for two-group classification such as bankruptcy prediction because the output variable ranges from 0 to 1 (Zhang *et al.*, 1999; Jo *et al.*, 1997).

The Gamma Memory Neural Network

A network with a memory is required to model temporal patterns. The GMNN is characterized by an adaptable memory, so it is ideal for modeling a situation with a multidimensional time-series, which is desirable in financial health applications. The recognition of temporal patterns requires two different memory structures in the neural network architecture. The *short term memory* decomposes the input signal into a fixed window of K parallel information streams and the *long term memory* is the traditional MLPNN static mapping from the time-series representation to a predicted output value. The short term memory captures the time-series structure of the data in a low dimensional space that preserves the dynamic properties of the system. Two important properties of the short term memory are the memory depth, D , (i.e., how far back in the time-series is information available) and memory resolution (i.e., the precision of the representation of these time-series values). A general representation for short term memory is the Gamma memory of De Vries and Principe, (1992). The Gamma memory is a short term adaptive memory formed by the temporal convolution of the time-series input signal, x_j , and a kernel function, w_{ij} defined in equation 1 (De Vries & Principe (1992)).

$$net_i(t) = \sum_j \sum_{n=0}^t w_{ij}(t-n)x_j(n) \quad (1)$$

De Vries & Principe (1992) show that the convolution model can be reduced to a system of differential equations with constant coefficients if the kernel function $w(t)$ is a solution of the following recurrence relation.

$$\frac{d^K w}{dt^K}(t) = \sum_{k=0}^{K-1} a_k \frac{d^k w}{dt^k}(t) \quad (2)$$

Solutions to the recurrence relation in equation 2 can be written as a linear combination of the functions $t^{k_i} e^{\mu_i t}$ where the μ_i are the eigenvalues of the system and the $t^{k_i} e^{\mu_i t}$ are the eigenfunctions. With this constraint, the short term memory structure is defined with a fixed number of kernels. To maintain stability, De Vries & Principe (1992) define a subset of functions from equation 2 that utilizes locally recursive structures. They define this subset as a *gamma memory* neural network consisting of gamma kernels. The gamma kernel is the integrand of the normalized gamma function defined as follows for discrete time-series.

$$g_k^\mu(t) \equiv \frac{\mu^k}{(k-1)!} t^{k-1} e^{-\mu t}, k = 1, \dots, K, \mu > 0 \quad (3)$$

For this memory structure, the kernel function is constrained as follows.

$$w(t) = \sum_{k=1}^K w_k g_k^\mu(t) \quad (4)$$

De Vries & Principe (1992) prove the approximation error of $w(t)$ by a linear combination of gamma kernels can be made arbitrarily small and that the gamma memory is stable for $0 < \mu < 2$. Details of the focused gamma neural network are given in De Vries & Principe (1992). An

advantage of the Gamma memory is that the parameter, μ , that controls memory depth and resolution can be adapted from the data during network training and does not have to be specified *a priori* by the user. This ability means that the short term memory depth and resolution will be determined empirically from the training data and the memory properties can evolve over time with changing system dynamics.

EXPERIMENTAL DESIGN

The MDA model, the Logit model, and the MLP neural network model are all static models that are fitted to the data at a single point in time. The financial ratio data for the point in time selected in this study is one year prior to bankruptcy. The MDA model and the Logit model will be modeled for the complete dataset for this specific point in time (i.e., 1 year prior to bankruptcy). Appropriate statistical tests will be run to assess compliance with underlying parametric model assumptions. However, given the nonparametric nature of the MLP neural network, a cross-validation approach will be needed to assess the errors in each model. A 9-fold cross-validation scheme will be used. Data for the 45 firms will be partitioned into nine groups. The same ratio of non-bankrupt to bankrupt firms will be used in each of the nine groups as are found in the overall dataset, creating a stratified sample for each group. Eight of the groups will be used for training the neural network and the ninth group will be used for testing. The groups will be rotated, with a different one serving as the ninth test group each time, testing all nine of the partitions.

The structure of the MLP neural network models considered in this study will consist of five input nodes, one for each of the five financial ratios, one hidden layer with varying numbers of nodes, and an output node. Four separate hidden-layer architectures will be tested for each partition. Rules of thumb suggested for determining the number of hidden nodes include: $n/2$, n , $n+1$ and $2n+1$, where n is the number of input nodes. Given the five input nodes, the range of possible hidden nodes is about 3 through 11. Consequently, we will run configurations with 4, 6, 8, and 10 hidden nodes each. The MLP neural network will employ the sigmoid activation function for all cases in this study. Given its nonparametric nature, the GMNN will follow the same 9-fold cross-validation scheme as the MLP in order to assess errors. Although the process for this network is different than the MLP, the GMNN will also be run for the same four hidden-layer architectures of 4, 6, 8, and 10 nodes. The activation function to be used for the GMNN is the hyperbolic tangent in all cases. The best MLP and GMNN models will be chosen from the experiment to be compared to each other and to the results obtained from the MDA and the Logit models. The McNemar's Chi Square test for statistical difference will be employed in comparing the predictive performance of each model (Siegel, 1956). This test is used in the two sample case, when samples are related. It is a nonparametric classification test for significance of changes, that is, differences that may occur before and after a particular treatment.

RESULTS AND ANALYSIS

As designed, the cross-validation scheme was run with nine partitions. Eight datasets of 40 observations each were used for training. One dataset of 5 observations was held back for testing. All model types were run. Since they are static models, MDA, Logit and MLP were run only for data from the last quarter prior to bankruptcy. The GMNN was run on all eight quarters

prior to bankruptcy to model the time-series. The GMNN and MLP architectures were limited to the case of 4 nodes in the hidden layer; time did not allow for running other hidden layer configurations. Results comparing the model approaches are given in Table 1. The results are significantly better for Logit and MLP as compared to MDA and the GMNN. Logit and MLP performed nearly identical in their misclassification (4/5%), while MDA and GMNN both performed poorly with similar misclassification rates of 13.3% and 15.6%, respectively. McNemar’s test was not run since these results are fairly clear cut.

TABLE 1 – RESULTS OF PRELIMINARY CLASSIFICATION RUNS

Model	Misclassification Number by Partition									Total	Percent
	1	2	3	4	5	6	7	8	9		
MDA	1	1	1	2	0	2	0	0	0	6	13.3%
Logit	0	1	0	1	0	0	0	0	0	2	4.5%
MLP	0	1	0	1	0	0	0	0	0	2	4.5%
GMNN	0	0	0	2	0	0	0	2	3	7	15.6%

CONCLUSIONS

Logit and MLP did very well in classifying the bankrupt and non-bankrupt companies one year prior to bankruptcy. Given that the data are only one year prior to bankruptcy, one would expect patterns to be easily represented by a good modeling approach. Also, their nearly identical performance is not surprising since Logit and MLP are very similar in their modeling approaches; Logit is essentially an MLP with no hidden layer. In previous studies, parametric models have been shown to perform comparably to neural networks such as the MLP when using data only one year prior to bankruptcy. For data two years or more prior to bankruptcy, the results of prior studies are less conclusive.

MDA’s poor performance compared to Logit and MLP is not surprising. This is generally consistent with the literature. However, the GMNN’s poor performance is not consistent with prior research. We believe that the poor performance of the GMNN is due mainly to improper data presentation to the GMNN software, which processes the time series data in parallel streams as a moving window of fixed size. Our dataset lists company ratios in a descending column with 8 rows for each company; each of the 8 rows represents a quarter of data. We believe in some cases the model convoluted ratio data from more than one company. Future research will focus on reformatting the data for proper entry. This should significantly improve the results for the GMNN model. In addition, there was time to run only one hidden-layer architecture for GMNN. Other architectures may also improve the performance of the GMNN.

REFERENCES

References available upon request from Scott Dellana at dellanas@ecu.edu

LESSONS LEARNED FROM OUTSOURCING PRACTICE IN A SME

*Mohammed H. A. Tafti, Department of Information Technology and Quantitative Methods,
Hofstra University, Hempstead, NY, (516) 463-5720, Mohammed.H.Tafti@hofstra.edu

Deb Sledgianowski, Department of Information Technology and Quantitative Methods,
Hofstra University, Hempstead, NY, (516) 463-4759, Deb.Sledgianowski@hofstra.edu

Jim Kierstead, Kanebridge Corporation, Oakland, NJ, (201)337-3200,
JKierstead@Kanebridge.com

ABSTRACT

Outsourcing IT functions has become a common practice for both large organizations and small-to-medium-sized enterprises (SMEs) globally. This paper discusses lessons learned in a recent ERP outsourcing effort from a practitioner's perspective within the context of existing research. The paper presents an analysis of the SME outsourcing practice, and discusses a case study of one company's experience concerning eight aspects of their outsourcing efforts. Their experience provides useful insights that are applicable to other organizations.

Keywords: outsourcing, offshoring, enterprise resource planning, small-to-medium enterprise

INTRODUCTION

Outsourcing, as a major IT and business strategy, has become a commonplace practice in recent years (Fish and Seydel, 2006; Gonzalez et al, 2006). According to Gartner, approximately 90% of all new businesses in the United States are SMEs and the trend is toward their increasing use of outsourcing (Mears, 2006). Offshore outsourcing (offshoring) of software development is when a company procures programming or other development processes from a supplier providing human resources that are located in a foreign country. Skills that are most likely to be offshored are those that don't require close interaction between the client and offshore service supplier or sharing of a common knowledge base with the client's business users; such as programming (Bullen et al, 2007).

Organizations generally expect to reduce their overall IT costs, focus on their core competencies, and gain superior technical resources when they outsource part or all of their IT functions. 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 (Fish and Seydel, 2006; Lee, 2004). However, without a careful consideration of the issues related to the enterprise-wide IT activities, any gain in the above areas can be offset by a significant loss of control, IT expertise, and overall corporate IT capabilities. Some of the risks associated with outsourcing include loss of IT expertise, hidden costs, inappropriate contract, limitation of

control and loss of privacy and security (Tafti, 2005). Given the lack of resources that SMEs have relative to larger organizations, offshoring can be a major undertaking for these firms (Carmel and Nicholson, 2005).

Literature has become more abundant on the subject of outsourcing with various viewpoints from advantages and potential risks to cultural, diversity, and ethical issues. While several papers have shed some light on the real-world cases of outsourcing practice (Barthelemy and Geyer, 2004; Chou et al, 2006; Rottman and Lacity, 2006), most articles on the subject of outsourcing are, however, theoretical and/or prescriptive. Further research is needed to provide more insights into bridging the gap between theory-based studies or prescriptive guidelines for outsourcing and actual outsourcing practice.

A purpose of this paper is to present an analysis of various aspects of a recent ERP outsourcing experience in Kanebridge, an SME in the industrial fastener industry. Eight aspects of outsourcing as practiced in Kanebridge and as discussed in the literature are presented in this paper which provide useful lessons learned from outsourcing practices. The next section presents a brief discussion of the company background and their outsourcing efforts. Eight outsourcing practices are identified and presented in this section. Then a comparative analysis is presented to illustrate these eight areas as practiced in Kanebridge and as addressed in the literature. Implications of the lessons learned are discussed in the conclusion.

BACKGROUND

Kanebridge Corporation is a leading wholesale distributor of industrial fasteners (nuts, bolts, screws, etc.). It is a privately-held corporation with 70 employees headquartered in Oakland, NJ with branches in California and Illinois. Their current annual sales exceed \$13.5 million (the company requested to withhold specific financial data).

Kanebridge's IT history included 30 years of in-house custom software development environments. Their legacy ERP system was developed before web-based systems became commonplace and practical. Due to "sunset" announcements by the vendors of their hardware and development environment, the opportunity arose for Kanebridge to consider replacing their custom ERP system with a purchased ERP package. At the low end of the price range Kanebridge could expect automated distribution and financial processes, but any other functionality is considered an "add-on" module with add-on costs. Implementation can be both costly and risky. Some of the questions Kanebridge considered were:

1. How much will they have to change the way they conduct their present business processes to accommodate a new ERP software package?
2. How easy is it to write custom code for the package and what is the difficulty in installing vendor updates to packages that have been customized?
3. What software development environment does the software package use and is it proprietary? For example, SAP uses a proprietary language called ABAP which requires skill in that particular programming language and their configuration tools in order to customize the code.

In 2003 Kanebridge made the decision to rewrite their applications using Java, instead of implementing a packaged software solution. Java's open source and cross-platform capabilities contributed greatly to their decision. They decided to consider using outsourced help in the software development process, mainly because they wanted to get the project done in a reasonable amount of time, and their IT manager was the only Kanebridge employee who had the necessary skills needed for the new software development environment.

Kanebridge's president first met the president of SourceCode Inc., a software consulting company, at a meeting of small business owners. The two company presidents became acquaintances and they informally talked about Kanebridge's legacy ERP system conversion project. SourceCode's president suggested that his company could provide Kanebridge with contract program developers to help them complete the project in a timely fashion, at a third of the cost to develop it onshore.

SourceCode submitted a bid for the project based on lines of code, number of screens, and other calculations. The two companies could not agree on a fixed price fee structure, but they finally did agree on a time-and-material fee structure. A time-and-material fee structure is considered more risky for the client since it puts the responsibility to expedite the work on the client company. A fixed price fee structure puts the responsibility to expedite on the outsourcing vendor, since they have incentive to finish the project as soon as possible, rather than collecting billable hours as with a time-and-material based fee structure. The time-and-material fee structure worked well for Kanebridge because it motivated them to look for productivity tools to save time and money. This resulted in the development of a custom software tool to convert Powerhouse code into Java code. Kanebridge rewarded the outsourced programmer who developed the conversion tool with a bonus payment.

Interestingly, Kanebridge and SourceCode did not use a formal contract. Rather, their agreement was based on a handshake. This was due in part to the pre-existing relationship between the two companies and the element of trust that already existed. Another reason was that there is an inherent motivation for both parties to keep their end of the agreement because SourceCode was getting their billable hours and Kanebridge was getting seasoned Java programmers who had ongoing experience working on the Kanebridge project; a win-win situation. One critical issue to Kanebridge, though, was that they had a legally binding agreement with SourceCode so that the source code developed during the project engagement belongs to Kanebridge and each SourceCode developer involved with the project signed a non-disclosure agreement.

As the project progressed, Kanebridge was ready to utilize additional outsourced Java developers. Two additional developers from SourceCode were added to the project to work at Kanebridge's NJ location, for a total of three onshore consultants. SourceCode recommended that Kanebridge use their offshore services for additional developers. They decided to "start small" and added developers in pairs as needed until there was a total of six offshore developers working in India.

The SourceCode onshore project leader handled most communication with the India project team. The offshore project team worked in two groups of 3 members each. Each group worked on a different module of the system. The project team was able to gain productivity using a "follow-the-sun" strategy, leveraging the 9-1/2 hour time difference by sending programming

specifications at the end of the onshore work day and receiving the code back from India the next morning. While Kanebridge was able to take advantage of the time difference to increase productivity, offshore productivity levels were two-thirds that of onshore. But this was more than offset by the hourly wage of the offshore workers relative to the onshore workers, which was one-third the cost. While cost savings may be the primary reason for U.S. firms to outsource (Rottman and Lacity, 2006), in the case of Kanebridge it was secondary to their need for a specific skill-set; although cost savings was the primary benefit of their offshoring endeavor.

Jennex and Adelakun (2003) suggest that traditional offshore application development tend to be highly structured with little or no changes to the requirements specifications, with minimal interaction and project management from the client. This was the case with Kanebridge and SourceCode; traditional offshore programming assignments were tasked to the India project team. When the programs were received from India and ready for testing, the testing took place onshore and any coding issues to be resolved as a result of the testing took place in NJ. The code did not go back to India to be fixed. This was done so that problems could be fixed right away and that the users could continue testing without too much lapse in time. Figure 1 (*not included in this summarized paper*) summarizes the above discussions showing eight characteristics regarding the outsourcing relationship between Kanebridge and SourceCode.

LESSONS LEARNED

Kanebridge's decision to develop their ERP system rather than buy a packaged solution and their ensuing outsourcing and offshoring experience should prove useful to IT and business leaders of other SMEs as they look to leverage information technology to enable and drive their business strategies. The decision-making and lessons learned described in this section provide insight for organizations to consider while implementing software development projects. The following discussion provides lessons learned based on our observations of eight outsourcing characteristics experienced by Kanebridge, and relates them to some of the relevant discussions in outsourcing literature. A summary of the characteristics identified, compared to the outsourcing literature, is presented in Table I (*not included in this summarized paper*).

Client-Vendor Relationship: This case demonstrated that networking events for small business owners would facilitate successful business relationships. This is consistent with the observations by Davis, Renzulli, & Aldrich (2006) in their study of voluntary association memberships. In Kanebridge's case, constructive interactions during networking events (Kern and Willcocks, 2002) have led to a successful outsourcing deal with SourceCode Inc. Furthermore, Kanebridge and SourceCode view each other as a "long time strategic partner". This "alliance" form of client-vendor relationship is a mature stage of outsourcing relationships (Gottschalk and Solli-Saether, 2006).

Company Size: Kanebridge's IT Director believes that one reason their outsourcing relationship worked so well is because the two companies are similar in size. He prefers to work with vendors similar in size because they often provide better customer service than the larger vendors, who may give preference to more lucrative contracts. While no empirical research to date has conclusively established a significant advantage from outsourcing client-vendor size parity, the prescriptive literature suggests that SMEs tend to use vendor companies of a similar size (Fox,

2006; Kobayashi-Hillary, 2006). Some, however, argue that this is because SMEs have difficulties in reaching out to large IT outsourcers (Paul, 2007).

Outsourcing Contract: Kanebridge has a very successful collaboration with their outsourcing vendor, SourceCode, without the use of a formal contract. This is not a recommended best practice particularly for large outsourcing deals. The outsourcing literature generally considers a detailed and comprehensive contract as a requisite condition for outsourcing success (Bryson and Sullivan, 2003; Jennex and Adalakun, 2003). However, there exists some empirical evidence supporting the notion that good-faith partnership and mutual trust, as practiced by Kanebridge, are critical to outsourcing success (Koh et al, 2004). 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” (Chou et al, 2006). A level of trust 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 their side of the agreement. The fact that there was a pre-existing relationship between the principles of the two companies created a trusting environment where Kanebridge felt comfortable that neither side would break their implied agreement.

Intellectual Property: According to the transaction-cost theory, the collaboration between a client and its outsourcing partner creates a value (or asset) the ownership of which must be agreed upon by both sides (Aubert et al, 2004). In the area of IT outsourcing, software development results in creation of source code that is considered as a significant intellectual property. Therefore, it is very important that the IT outsourcing contracts specify ownership of this intellectual property (Walden, 2005). Although no general outsourcing contract has been signed between the two companies as discussed above, Kanebridge insisted on, and got a signed document assigning the intellectual property right of the resulting source code to Kanebridge.

Custom functionality: The common wisdom in software development is to develop a solid business case for *not* implementing a commercially available software package (Keil and Tiwana, 2006; Lucas 2005). If an existing package will solve your business problem, then it can be a much faster and cheaper solution than insourcing or outsourcing application development. This general guideline is even more relevant in the case of ERP implementation particularly in an SME environment because the cost of the customization process is generally higher than the cost of the ERP acquisition (Chen and Soliman, 2002; Quiescenti et al, 2006). In the case of Kanebridge, once they realized no existing ERP package would provide their unique custom functionality, which is a core competency, they stopped analyzing alternative solutions. They wanted the flexibility to be able to easily modify the code as their business processing needs evolve.

Task Complexity: One of the main determinants of IT sourcing strategy is the degree of task complexity. Organizations tend to choose more structured and less complex applications to outsource and keep more complex ones for in-house development (Jennex and Adalakun, 2003, Kishore et al, 2004/2005). The ERP components chosen by Kanebridge are considered to be highly structured which required minor changes to specifications. As the organizations that engage in outsourcing deals gain further experience, more complex projects that require greater

client contact and project management are selected to be outsourced (Mirani, 2006; Xue et al, 2004/2005).

Communication: In their outsourcing deal, Kanebridge required the vendor's project leader to learn and fully understand the company's outsourcing needs and objectives. SourceCode provided an onshore contact point in Kanebridge's New Jersey office to ensure the availability of someone who understood the offshore culture and who could effectively facilitate interface among diverse backgrounds of offshore developers. This is highly consistent with research findings in the outsourcing literature regarding the positive impact of information exchange and clear communication of diverse knowledge on the level of team performance (Koong et al, 2007; Ting-Peng et al, 2007).

Vendor's Skills: One of the main reasons of IT outsourcing is to take advantage of the vendor's technical skills (Aubert et al, 2004). Kanebridge's expectations from their outsourcing vendor included not only a proper level of technical skills, mainly java programming, but also a clear understanding of their ERP project objectives. It was critical that the outsourced IT staff were experienced Java programmers proficient with the software development products being used. Having a project contact point that was knowledgeable of Kanebridge's ERP needs and the technology being used, and could successfully communicate project requirements with the offshore project team was a critical factor leading to a very successful outsourcing venture.

CONCLUSION

This case represents a successful endeavor by a medium-sized business to effectively manage their information technology resources to efficiently solve a business problem and in doing so, preserving an information system that provides them with a competitive advantage while establishing a new software development environment. In the process, they learned how outsourcing can be used as a tool to provide selected resources as needed. The practices applied and decisions made by the case company provide insight for organizations to consider while implementing software development projects.

This case demonstrated that networking events for small business owners can work to facilitate business relationships, as suggested by Davis, Renzulli, & Aldrich (2006). A networking event provided an informal opportunity for Kanebridge to connect with SourceCode Inc., their software service provider. One of the reasons why their outsourcing relationship worked so well is because the two companies are both similar in size. Kanebridge's IT Director prefers to work with vendors similar in size because they often provide better customer service than the larger vendors, who may give preference to more lucrative contracts. And surprisingly, Kanebridge was able to have a successful agreement with SourceCode without the use of a formal contract. This demonstrates that partnership and trust are critical to outsourcing success. A level of trust can reduce the transaction costs inherent in a relationship between two parties. While a detailed, comprehensive formal contract is highly recommended throughout the literature, having a close and trust-based relationship can act as a proxy to formal contracts, and could lower the risks that precipitate the need for outsourcing contracts.

**NOTE: Complete paper including figures, tables, and references is available upon request.*

IMPACT OF TRUST, INTERDEPENDENCE AND PERCEPTIONS OF STRATEGIC SALIENCE ON IT OUTSOURCING SUCCESS

Narasimha Paravastu, Metropolitan State University, Minneapolis, MN, (612) 659 7288,
narasimha.paravastu@metrostate.edu

David Gefen, Drexel University, Philadelphia, PA, (215) 895 2148, gefend@drexel.edu

ABSTRACT

This research is about the factors of relationship management that contribute to the success of an IT Outsourcing project. This research tests the impact of strategic salience on success of an IT outsourcing project, and impact of trust, and interdependence of parties in a relationship on strategic salience. Data were collected from an online survey of 166 IT professionals all over the United States, and tested using structural equation modeling. All the hypotheses were supported. The implications of the results to theory and practice are discussed.

INTRODUCTION

IT Outsourcing is defined as handing over of all or a part of an organization's IT functions or services to an external service provider in order to achieve organizational goals (Cheon, Grover et al. 1995). The external service provider is called the vendor and the organization obtaining the service is known as the client. The IT functions that are commonly outsourced include systems operations, applications development, applications maintenance, help desk and end user support, and systems planning and management (Grover, Cheon et al. 1994; Dibbern, Goles et al. 2004). IT outsourcing affords several benefits to the clients such as the cost economies, increased profitability and reduction in costs, financial leverage, and access to expertise (Loh and Venkatraman 1992; Lacity and Hirschheim 1993; Lacity, Hirschheim et al. 1994). Companies that outsource their IT expect to achieve cost savings of 25 to 30 percent (Lacity, Hirschheim et al. 1994; Deloitte 2005; Gartner 2005). Because of the above reasons and because of the general economic factors, pressures for increased returns and enticing incentives offered by the vendors (DiamondCluster 2002) IT outsourcing has grown phenomenally. By 2007, global spending on IT outsourcing is expected to top \$50 billion a year (Gartner 2005). However, IT outsourcing contracts often run into problems. Market reports indicates that about 70% of the companies experienced negative outcomes in their outsourcing contracts (DiamondCluster 2002; Deloitte 2005) and about 80% of the contracts need re-negotiation (Deloitte 2005; Gartner 2005). Among the central reasons why IT outsourcing fails in the post implementation phase are executives failing to see the strategic salience of the project (Lacity and Hirschheim 1993; Lacity, Hirschheim et al. 1994; Barthelemy 2001) and inadequate post-contract phase relationships with the vendor (Lacity and Hirschheim 1993; Lacity, Hirschheim et al. 1994; Willcocks, Fitzgerald et al. 1996). Taking this observation a step further, therefore, implies the need to nurture executives' recognition of the salience of the IT outsourcing project and their relationship with the vendor well into the post implementation relationships. Based on previous research on trust in inter company contracts (Kumar 1996) including IT implementation (Gefen 2004), this study tests the relationship of strategic salience to success of an IT outsourcing

project, and impact of trust, and interdependence of parties in a relationship on strategic salience. Generally, past research has recognized the need to nurture trust and interdependence between the parties in the post implementation phase of outsourcing contracts (Sabherwal 1999; Barthelemy 2001; Kern and Willcocks 2001). This study takes a new stand on why this is so by introducing strategic salience and showing its effect over and above the contribution of trust and interdependence. Moreover, trust in IT outsourcing relationships has been tested in the past through interviews and qualitative techniques (e.g.: Sabherwal 1999) or by means of a single item question as a part of a relationship (e.g.: Lee and Kim 1999). Introducing this stand, the study examines trust in IT outsourcing relationships through a social exchange (Blau 1964) perspective and conceptualizes client's trust in vendor in accordance with the past research (Luhmann 1979; Mayer, Davis et al. 1995; McKnight, Choudhury et al. 2002; Gefen, Karahanna et al. 2003). Based on similar constructs in previous research, the study develops scales for measuring strategic salience in the context of IT outsourcing.

Outsourcing success is the ability of the client firm to achieve technological, financial, and strategic goals for the organization and a level of satisfaction with the vendor (Grover, Cheon et al. 1996; Lee and Kim 1999). Strategic salience is the perception of the managers involved that the collaboration with the IT outsourcing vendor will be beneficial to the organization and yield strategic advantage. Trust is the general level of the client's willingness to be vulnerable to the actions of the vendor based on their belief in the vendor's trustworthiness (Mayer, Davis et al. 1995; McKnight, Choudhury et al. 2002; Gefen, Karahanna et al. 2003). Interdependence is the clients' perception of a firm's dependency on its vendor firm, relative to the vendor firms' dependency on it (Lee and Kim 1999).

Theoretical Background and Research Model

The research model hypothesizes a positive relationship between trust and interdependence. Trust and interdependence are related to strategic salience. Strategic salience is hypothesized to be positively related to IT outsourcing success.

IT Outsourcing Success

The dependent variable in this study, IT Outsourcing success represents the objective of the IT outsourcing process from both the client side and the vendor perspectives. An IT outsourcing project can be said to be successful if the expectations of the organization underlying the outsourcing decision are met. The general organizational expectations behind outsourcing decisions include financial, strategic, technical, and relational expectations (Lacity, Hirschheim et al. 1994). Financial expectations include economies of scale of the vendor, control over IT costs, and restructuring IT budgets. Strategic expectations include refocusing on the core competencies. Technical expectations include access to new technology, and expertise better quality of service, eliminating IT incompatibilities at the time of mergers and acquisitions and sourcing IT from vendors in case of small startups (Lacity, Hirschheim et al. 1994). Based on the above, and as defined by previous research (Grover, Cheon et al. 1996; Lee and Kim 1999), outsourcing success is defined in this study as achieving technological, financial, and strategic goals for the organization. Strategic success is the firm's ability to focus on its core business; outsource routine IT activities so as to focus on strategic uses of IT; and enhance its IT competence (Loh and Venkatraman 1992; Loh and Venkatraman 1992; Grover, Cheon et al. 1996). Financial success is the firm's ability to utilize the economies of scale of the vendors and reduce costs, and achieve economies in profitability (Loh and Venkatraman 1992; Loh and

Venkatraman 1992; Grover, Cheon et al. 1996). Technological success is the ability of the firm to gain access to IT expertise and latest developments in IT and avoid IT obsolescence (Grover, Cheon et al. 1996).

Perceptions of Strategic Salience

Strategic salience is the client's perception that its collaboration with the outsourcing vendor is beneficial to the organization and yields a strategic advantage. The alignment of the IT function with the organizational strategy is recognized as a source of competitive advantage for the organization (Barney 1991; Kearns and Lederer 2003) and results in increased profitability (Sabherwal and King 1991; Henderson and Venkatraman 1999; Tallon, Kraemer et al. 2000). IT Outsourcing is a strategic decision (Grover and Teng 1993) for the organization which could potentially result in increased profitability and cost savings among other benefits (Loh and Venkatraman 1992; Lacity, Hirschheim et al. 1994). In order to realize the positive results of alignment of IT with business strategy, it is critically essential that the alignment processes result in increased understanding and participation of the IT managers who implement them. Past research found that IT projects were successful if the executives realized the importance of implementation of the projects and resulted in escalated costs, unrealized expectations and project failures when the executives failed to understand the nature of the IT function and importance of the IT project implementation (Lacity and Hirschheim 1993; Lacity, Willcocks et al. 1996). Therefore, the strategic salience of the IT outsourcing project, monitoring the project performance, and maintaining the client vendor relationships among the IT outsourcing executives is critical for the success of an IT outsourcing project.

Interdependence

Interdependent business relationships in general are mutually mutual beneficial, and the parties have considerable influence over the actions of the other party (Anderson and Narus 1990). This extends to the specific context of IT outsourcing also (Lee and Kim 1999). The benefits enjoyed out of the relationship include could include transaction cost benefits, higher economic returns, and better productivity for the client as well as the vendor (Kalwani and Narayandas 1995; Narayandas and Rangan 2004). Vendors depend on their clients for its business, and clients depend on the vendors for execution of the outsourcing project. From a client's point of view, interdependence is critically important because, clients often eliminate their in-house IT function when deciding to outsource. Such dependence on the vendor could either develop asymmetries of power favoring the vendor which contributes to opportunistic behavior of the vendor and project failures, or, could lead to a strategically salient and mutually beneficial relationship for both parties.

Trust

Trust is recognized as a significant factor in business relationships (Kumar 1996). This study adopts a social exchange (Blau 1964) theoretic approach to explore the influence of trust in IT outsourcing relationships. Social exchange defines trust as the willingness of a party to be vulnerable to the actions of another (Zand 1972; Mayer, Davis et al. 1995) based on the expectation that that the other party will not indulge in opportunistic behavior by taking advantage of the situation (Gefen 2002). This study adapts this conceptualization of trust because of its relevance to outsourcing relationships from a social exchange perspective. In the context of outsourcing trust in vendor is defined as the willingness to be vulnerable to the actions of the IT

outsourcing vendor, with an expectation that the vendor will not engage in opportunistic behavior. Trust in the vendor is the general level of trust in a specific IT outsourcing vendor, without reference to any particular behavior or characteristic of the trusted party (Chen and Dhillon 2003).

Trust reduces perception of risk in social interactions (Luhmann 1979) as well as in business relationships (Mayer, Davis et al. 1995; McKnight, Choudhury et al. 2002; Gefen, Karahanna et al. 2003). Trust is a key element in interpersonal relationships (Lewis and Weigert 1985) as well as in business relationships such as ecommerce transactions (McKnight and Chervany 2001; Ba and Pavlou 2002; Gefen, Karahanna et al. 2003) and client-vendor relationships in ERP (Gefen 2004). Trust strengthens interpersonal relationships by promoting cooperation between the parties in relationship (Gambetta 1988). Though some of the earlier studies (Lee and Kim 1999) have included trust as a part of the relationship itself, social exchange theory (Blau 1964) recognizes trust as a concept distinct from the relationship itself and as a key contributory factor to the relationship.

HYPOTHESES

Strategic Salience and Outsourcing Success

IT outsourcing is a strategic decision taken by management after evaluating various options of outsourcing and the organizational objectives (Loh and Venkatraman 1992; Grover and Teng 1993; Lacity, Hirschheim et al. 1994; Lacity and Willcocks 2001). Capturing the perceptions of these managers who oversee the project implementation, this recognition of the strategic salience of the IT outsourcing is closely related to their perception of IT outsourcing success. (Lacity, Willcocks et al. 1996). Based on the results from the past research it is hypothesized as follows:
H₁: Strategic Salience is positively related to outsourcing success.

Interdependence and Strategic Salience

Interdependence is defined as the client's perception of a firm's dependency on its vendor firm, relative to the vendor firms' dependency on it (Lee and Kim 1999). In business relationships where the parties in the relationship depend on each other, each party makes its own scrupulous estimate of the probable behavior of the other party. The parties continue the relationship based on the calculation that the other party will not violate the expectations because of the mutuality of dependence (Shapiro, Sheppard et al. 1992; Doney and Cannon 1997). Therefore, the relationship is perceived as beneficial and important to both parties (Paul and McDaniel 2004). Therefore it is hypothesized that:

H₂: Interdependence is positively related to perceptions of strategic salience.

Trust and Strategic Salience

Past research indicates that trust in a vendor is created partly by past experiences with the trusted party (Lewicki and Bunker 1995; Rousseau, Sitkin et al. 1998; Ba 2001), transference from a trusted source (Milliman and Fugate 1988; Stewart 2003), institutional guarantees (Pavlou and Gefen 2004), and the characteristics of the trustee such as vendor's ability, benevolence and integrity (Mayer, Davis et al. 1995; Gefen 2002; McKnight, Choudhury et al. 2002). Trustworthiness of the relational partners is considered a source of strategic advantage (Barney and Hansen 1994). Therefore, the clients trust in the vendor, based on multiple trust modes and

trustworthiness attributes, will increase the perceptions of strategic salience of the relationship with the vendor. Therefore, it is hypothesized that:

H₃: Trust is positively related to perceptions of strategic salience.

Trust and Interdependence

Trust is an intention to be vulnerable to the actions of another party in a relationship. The parties in a trusted relationship behave in a manner consistent with the expectations of the relationship (Mayer, Davis et al. 1995; McKnight, Choudhury et al. 2002). Interdependence is the observed behavior of the parties in the relationship where one party depends on the other. Based on this being an integral part of social exchange theory (Blau 1964), it is hypothesized that:

H₄: Trust is positively related to interdependence.

METHODOLOGY

The instrument for measuring the above constructs was developed using validated scales for measuring similar constructs in past research. Items measuring outsourcing success were adapted from Grover et al. (1996) and Lee et al. (1999). The items measuring Strategic Salience were created for the purpose of this research based on Kearns et al. (2003). Items for interdependence were adapted from Lee et al. (1999). Items measuring trust in vendor were adapted from Gefen, Karahanna and Straub (2003). All the questions were Lickert scales of seven (strongly agree to strongly disagree). Each question was reviewed by experts in the area to check for the accuracy of intent of the question, clarity, and consistency with the definition of the construct.

Before the survey was posted online for data responses from targeted sample, the instrument was subjected to extensive pretests to check for content and construct validity and reliabilities. The questionnaire was tested and evaluated for its online layout and functionality, clarity of question wording, for content and face validity of the new questionnaire items designed for the study by several Ph.D. students who were in various advanced stages of their dissertation. The questionnaire items designed for this study (interdependence and strategic salience) were rigorously pre-tested using Q sort and elimination techniques. The final set of questionnaire items that are used for analysis in this study, their factor loadings and construct reliability alpha met or exceeded the acceptable psychometric thresholds.

Sample and Data Collection

Data Collection, Analysis and Results

An online survey instrument with items measuring the constructs were administered to the targeted sample which consisted of professionals all over the United States involved with IT outsourcing. The data were analyzed with AMOS 6.0. Fit indices indicated a good fit (Wheaton, Mutchen et al. 1977; Hair, Anderson et al. 1998) between the model and the data with, $\chi^2_{87} = 159.31$, GFI=.89, AGFI=.86, NFI = 0.93, CFI = 0.98, and RMSEA = 0.071. Reliability coefficients, all within the acceptable range of 0.80 to 0.96. All the hypotheses were supported at the 0.01 level of significance or better. Perceptions of strategic salience was significantly related to outsourcing success at the 0.001 level, explaining 0.574 of the variance in outsourcing success. Interdependence was significantly related to perceptions of strategic salience at the 0.01 level and Trust was significant at the 0.001 level to perceptions of strategic salience.

Interdependence and trust together explained 0.432 of variance in perceptions of strategic salience.

Discussion

Important contribution to theory is that, this study operationalized and tested trust in vendor adapting the conceptualization of trust, from the past research and using scales that were validated in past research. The conceptualization and measurement of trust from a social exchange perspective helps to explain to some extent, conflicting findings from past research regarding interdependence of parties in a relationship which is explained in the following paragraph.

Several studies found that interdependence of the parties contributes to strengthen business relationships (Ring and Van de Ven 1992; Ganesan 1994; Ring and Van de Ven 1994; Klepper 1995; Gattiker and Goodhue 2005). However, in IT outsourcing relationships, Lee and Kim (1999) posited a positive relationship but found a significant negative impact of mutual dependence of parties and the quality of relationships, which was attributed to the client's perceptions of risk. The present study corroborates the findings in Lee and Kim study (Lee and Kim 1999). Trust reduces the perceptions of risk (Gefen, Karahanna et al. 2003), and therefore, in trusted relationships, interdependence is positively related to IT outsourcing success.

Managers' understanding (or lack it) of the nature and importance of IT outsourcing project has been recognized in past research as an important factor in the success or failure of an outsourcing project. However, this has not been empirically tested using validated scales. This study developed strategic salience construct to measure the managers' understanding of importance of the projects, and tested the items measuring the perceptions of strategic salience of the managers involved in IT outsourcing projects.

Implications for Practice

Relationship management and effective monitoring of the post-contract phase, is essential for success of the IT outsourcing projects. This research highlights some important considerations which have implications before and during the relationship phase in the IT outsourcing contract in the areas like vendor selection, monitoring and closely working with the vendor. Recognition that the relationship is based on interdependence which is mutually beneficial for the client as well as the vendor will help to increase the awareness about the nature of the relationship and importance of the project success for the parties. Some of these relational characters can be verified at the time of vendor selection, through extensive research into the characteristics of the vendor firms, and the compatibility of the client-vendor work cultures, and the history of other projects executed by the vendors.

Conclusion

This study is based on a sample of 166 IT outsourcing professionals all over the United States. The sample in this study is reasonably representative of the population as the sample included professionals in the area of IT projects, and IT outsourcing across a variety of industries. The strength of the study is that, the items used in the study to measure the constructs were drawn from the past research, and extensively pre-tested and validated. The study has strong internal validity as evidenced in the psychometric properties of the scales, factor loadings and the reliabilities of the items.

References available upon request from Narasimha Paravastu.

A CROSS-SECTION EXAMINATION OF KMS ON FIRM PERFORMANCE

Edward T. Chen, U. of Massachusetts Lowell, (978) 934-2756, edward_chen@uml.edu

ABSTRACT

Despite the anecdotal contributions of knowledge management systems (KMS) to firm performance, a cross-section examination of KMS effects is still lacking. This study compares KMS adopters with non-adopters based on a set of established financial ratios. The results indicate that the firms with KMS outperform their counterparts in many areas.

Keywords: Knowledge management systems, firm performance, technology adoption

INTRODUCTION

Over the last decade, despite the increasing anecdotal evidence that KMS contributes to the firm performance, the empirical results relating KMS investments to firm performance measures have been equivocal. However, plenteous studies have relied exclusively on traditional accounting-based measures of firm performance, tending to downplay the contributions of KMS to performance dimensions. A review of literature reveals a dearth of empirical research on how to successfully develop and implement KMS to enhance firm performance. Much of the existing KM research focuses either on the use of various technologies to acquire or store knowledge resources or on the conceptual nature of KM [1, 26, 33]. Whether adopting KMS can improve firm performance remains uninvestigated. Given that IS researchers and practitioners often debate over the contribution of IS investment to firm performance, this study accentuates the impacts of implementing KMS on firm performance.

The benefits of adopting KMS are widely claimed even though the KMS wave has hardly touched the pragmatist shore. The main question this study tries to answer is: "How justifiable are these claims of KMS benefits?" Thereby, this study attempts to analyze the financial statements of KMS adopters who expect competitive advantage to show up in accounting performance measures. The purpose of this study will be accomplished in the following steps. The first step is to identify KMS adopters and to measure their performance relative to a matched control group. The second step is to measure the performance of KMS adopters relative to a matched control group using DuPont Analysis.

Knowledge management systems (KMS) are IT-based systems developed to support and enhance the organizational processes of knowledge creation, storage, retrieval, transfer, and application [1]. Not all KM initiatives involve an implementation of IT, but many KM initiatives rely on IT as an important enabler [12]. The objective of KMS is to reveal three common applications: the coding and sharing of best practices; the creation of corporate knowledge directories; and the creation of knowledge networks [1]. Critical to KMS, information technologies, customer relationship management (CRM) and business intelligence (BI) for example, assist in speeding up communications, eliciting tacit knowledge, and improving customer service [23]. There is no single role of IT in knowledge management just as there is no single technology comprising KMS. The potential benefits of KMS include productivity and quality improvements in key areas such as product reliability, customer service, and productivity. KMS is thus expected to enhance firm performance through efficiency and effectiveness gains.

IMPACT OF KMS ON FIRM PERFORMANCE

This study is based on the argument that KMS play a central role in the creation and realization of knowledge-based synergies across different units of KMS adopters. Research indicates that a successful adoption of IS to support business strategies can help the firms gain a superior financial performance [5, 18, 19, 22, 24]. KMS, as a class of specific IS applied to manage organizational knowledge, can facilitate the efficient and effective sharing of a firm's intellectual resources [31]. The fundamental premise is that firm performance improved by KMS may produce long-term sustainable competitive advantage [1, 36, 39]. KMS adopters believe that the value of KMS will exceed its cost. From the knowledge-based view, firms can devise strategies to create and sustain advantages from investments in IT [14]. Researchers have shown that a firm's ability to effectively leverage its IT investments by developing the strong IT capability can result in a superior firm performance [5, 34]. Based on the knowledge-based view, this study hypothesizes that KMS adopters could add more business values and decrease costs by creating knowledge-based synergies than non-adopters. KMS could enhance KM capability and performance of the firm by facilitating exchange of knowledge resources across the functional units of firms.

HYPOTHESES DEVELOPMENT

This study examines the firm performance attained by KMS-adopting firms. It is expected that, other things being equal, KMS adopters will outperform competing firms that are not adopting

KMS. This insight necessitates comparing the differential returns of the adopting firms with those of competing firms. Differential measures have been used in former studies to examine the long-term performance pertaining to competitors [5, 34]. The first hypothesis of the research is:

H1: Firms that adopt KMS will show greater performance as measured by financial performance ratios than that of non-adopters with the same firm size and in the same industry.

A firm's value of creating historical performance and future potential can be monitored through use of the DuPont Analysis. Developed by scientists at DuPont about a century ago to track the firm's performance in its diversified investments, this analysis looks at net profit margin (NPM) and asset turnover (ATO) as the building blocks to return on assets. The DuPont Analysis can help to identify and quantify value drivers and ultimately develop strategies to improve ROA and create value [16]. Increased efficiency is likely to show up in efficiency measures such as inventory turnover and total asset turnover. In addition, increased efficiency will appear in profitability measures such as gross profit margin and net profit margin to the extent that fixed costs are a component of the cost of produced goods [13]. Fairfield and Yohn [17] provide evidence that disaggregating ROA into ATO and NPM does not provide incremental information for forecasting the change in ROA one year ahead, but that disaggregating the change in ROA into the change in ATO and the change in NPM is useful in forecasting the change in return on assets one year ahead. DuPont Analysis makes possible a simultaneous analysis of efficiency and profitability, and it shows how the NPO and ATO interact to determine ROA.

A practical reason for employing ROA to evaluate IT investment is the extent of ROA as a measure of competitive advantage. ROA measures the accounting income return to a company for each dollar of assets employed in the business. As shown in equation (1), this can be expressed as a mathematical formula, consisting of a profitability measure (NPM) and efficiency measure (ATO). NPM measures income from ongoing operations per dollar of sales, while ATO measures how many dollars in sales of the firm is able to produce for each dollar invested in total assets, or in other words, how efficiently management utilized assets to generate sales [13]. This study proposes the use of the DuPont framework to address the following research question: Where do we expect to observe the performance payoff for firms with a competitive advantage from adopting KMS? The answer to this question will come from the analysis of the differences in ROA, NPM, and ATO between KMS adopters and non-adopters.

$$\text{Return on Assets (ROA)} = \text{Net Profit Margin (NPM)} \times \text{Total Asset Turnover (ATO)} \quad (1)$$

Based on the above discussion, this study proposes our second hypotheses:

H2: Firms that adopt KMS will improve profitability and efficiency in firm performance than non-adopters with the same firm size and in the same industry.

RESEARCH METHODOLOGY

This study investigates the firm performance of KMS based on the announcements of KMS vendors. While no evidence indicates that the market reaction to IS announcements will differ based on the vendors, there is corollary evidence that the market reacts more positively and strongly to higher quality auditors [19, 37]. Financial statements provide information that is useful to business decisions. Using the archival financial data extracted from the COMPUSTAT includes the financial statements of publicly-traded companies in the United States. To identify firms that adopting KMS, this study carefully selects the qualified KMS vendors from the *KM World Magazine*, which lists the KM World's 100 companies and provides different kinds of KM applications in its March 2003 issue. This study searches Reuters.com for the key developments of KMS vendors publicly disclosing adopters that acquired their KM application. Since Reuters.com provides the data after 1999, with this constraint, this study taps into Lexis-Nexis and searches for KMS adopters to widen the sample size. This study defines the event as a public announcement of a firm's KM effort. The online search features of Lexis-Nexis are used to search five leading news sources-*Business Wire*, *PR Newswire*, *The New York Times*, *The San Francisco Chronicle*, and *USA Today*-for announcements containing the words: "knowledge management", "document management", "customer relationship management," "business intelligence", and "knowledge management system" in headline and lead paragraphs.

KMS adopters are ascertained from Reuter.com and Lexis-Nexis Academic Universe's News Wire Service Reports. Firms publicly announcing a KMS implementation from January 1, 1995 through October 31, 2004 are initially included in the sample. A keyword search is employed using a combination of the searched terms with the name of each of the following KMS vendors: Interwoven, IBM, Primus, Oracle and so forth. This initial search reveals 6,536 hits. Each announcement is carefully examined for the validity. The examination process filtered out general business announcements such as general trends in KM, a partnership between two or more KM vendors, or the sales or performance of a specific KM product from a vendor's perspective. The first cut turned out 673 cases.

Since the COMPUSTAT only lists publicly-traded companies in the United States, this study

excludes non-profit organizations, government agencies, private companies, and KM vendors from these 673 cases. The valid sample of 270 KMS adopters has then been identified by further excluding confounding announcements such as merger, acquisition, and spin-off. This study follows the suggestions of former research [8, 9, 10] and does not count financial data of the firm in the immediate following fiscal year when KMS adopters were announced. Though the sample size will be further reduced, this rigorous approach allows this study to examine the firm performance objectively. To accommodate three-year comparison, this study excludes the firms that adopt KMS after 2001. Finally, 103 KMS adopters are selected for testing proposed hypotheses. KMS adopters have to satisfy the following criteria to be included in the samples of this study. First, the KMS adopters have to be listed on the COMPUSTAT. Additionally, adopters have to be active as of the end of 2004 fiscal year.

RESULTS AND DISCUSSION

Results of the regression analysis that seek to examine the effects of KM capability between KMS adopters and non-adopters. It is important to note that in analyzing the effects of KMS, a significant positive coefficient for the dummy variable relating to profit ratios and a significant negative coefficient for the dummy variable relating to cost ratios. In the cost ratios, the results indicate that selling and general administration expenses divided by sales (SGAS) are significantly different between KMS adopters and non-adopters in all three consecutive years. The results of the number of employees divided by sales (EMPREV) and the operating expense divided by sales (OEXPS) are significantly different between KMS adopters and non-adopters in the first year, but not significant in the second and third years. In the profit ratios, ROA of adopters differs apparently from non-adopters in all three consecutive years. The result of ROS is significantly different between KMS adopters and non-adopters in first and second years, but is insignificant in the third year. Operating income divided by asset (OPINA) is significantly different between KMS adopters and non-adopters in the third year, but is insignificant in the first and second years. Overall, the test results partially support the hypothesis 1.

The results of DuPont Analysis, a method that helps to investigate possible competitive advantages achieved through the successful adoption of KMS in efficiency (i.e., asset turnover, ATO) and profitability (i.e., net profit margin, NPM). Both ATO and NPM show no significant differences either before KMS adoption or during the year when KMS adoption is announced. These results provide a fair ground for advantage comparisons after KMS adoption. After all three years of adopting KMS (t_{+1} , t_{+2} , t_{+3}), the advantage lies in both asset utilization and net profit margin, except the net profit margin in the second year, are significant and that asset

utilization contributes to the advantages. As a result, hypothesis 2 is strongly supported. The implication from this result is that the most often cited advantage of IS, increased efficiency, can lead to a competitive advantage, but increased profitability is equally important.

CONCLUSION

Using a matched pair design to compare KMS adopters and non-adopters, this study examines the longitudinal impact of adopting KMS on firm performance. According to the statistical results, major findings are revealed that in the cost ratios. The results indicate a significant difference of administrative expenses (SGAS) in the consecutive years after implementing KMS. The operating expenses to sales (OEXPS) and the number of employee to sales (EMPREV) are significantly different in the first year between KMS adopters and non-adopters. As the main purposes of KMS are to reduce administrative expenses and to improve productivity by maximizing KM capability. Providing knowledge workers necessary access to the knowledge repository allows them to efficiently review and effectively retrieve the timely information for better decision-makings. This study supports the claim that firms adopting KMS does help lessen the administrative expenses and strengthen productivity.

In the profit ratios, the results show the significant difference between the return on assets (ROA) of KMS adopters differed apparently from non-adopters in all three years. The results of return on sales (ROS) are significantly different between KMS adopter and non-adopters in the first and second years. The operating income to assets (OPINA) is significantly different between KMS adopter and non-adopters in the third year. These results show that KMS adopters have superior KM capability than non-adopters in light of ROA, ROS and OPINA that determine the profitability of firms.

The results of the DuPont Analysis point to the competitive advantages in both efficiency and profitability through a successful adoption of KMS. The results indicate that both asset utilization and net profit margin are significant. The implication of this result is that the increased efficiency, the most often cited advantage of IS, can lead KMS adopters to a competitive advantage. Nonetheless, the increased profitability remains important. The findings thus provide new insights into the productivity paradox associated with KMS adopters and confirm that KMS adoption helps firms gain a competitive advantage over non-adopters.

“References available upon request from author.”

A NOVEL APPROACH TO MODELING A HIGHLY DYNAMIC SYSTEM: IS SECURITY

Najib Saylani
Hofstra University
(516)463-5716
Najib.saylani@hofstra.edu

ABSTRACT

When dealing with the issues of security and privacy, information systems can be considered examples of highly dynamic systems especially when you factor in the human day to day interactions with the system. The IS and its supporting IT infrastructure can be considered a complex system due to the many actions and reactions initiated within and from outside. In the context of security, activities that may compromise the security of an IS could be traced, monitored, and, when breaches occur, an efficient set of forensic processes could be conducted. Patterns of such compromises can go unnoticed for a long time and this may be due to the existence of transparent relationships between all sub-entities comprising an IS. We propose an approach using Cellular Automata (CA) to model such complexity for such dynamic system

CELLULAR AUTOMATA BACKGROUND

Cellular Automata were first used by *J. Von Neumann*. His idea was to construct a reliable system from unreliable parts. The subject developed to more sophisticated processes like the game of Life by *John Conway*. Cellular Automata are discrete space-time models. A set of units (cells) arranged in the Euclidian space with one or two dimensions spaces. The cells are also called sites. According to some local rules (local to a given site), a cell will take a certain value (i.e. binary value 0 or 1) at a given time t , a value which depends on the value of neighboring cells. We talk about left/ right or north/south effects or even diagonal effects. The state of a given cell changes from time t to time $t + 1$ according to what happens to its adjacent cells. In general, CA are dynamical systems within discrete space and time. The set of cells is a regular lattice and the states of the cells are synchronously updated according to a deterministic set of local rules. Cellular Automata techniques are applied to many fields such as biology to study clustering of cells and clustering of cell activity; CA were use to study crystal growth and in physics, they are used to study the interaction of different particles. These uses are of particular importance to the study described herein. [2]

CA'S RELATION TO THE MODEL

In this work, and in the context of security and the human factor, we should note that the study of patterns that are representative of security concerns can be modeled as a set of scattering phenomena between different internal and external agents affecting the IS. With the proposed techniques, the determination of future and past behavior of patterns scattering processes should be possible. The proposed system enforces the idea of the need of a system capable of not only highlighting the process of patterns scattering but also reconstructing all traces of the system behavior by considering the simple manifestation state of an individual entity in relationship with all other sub-entities. The IS system in the context of security is an ever evolving dynamic system (a dynamical one of course). To expand on the use of CA in this research, another category of CA will be explored:

Invertible Cellular Automata (ICA), offer a way of not only investigating the consequences of patterns scattering in the system but also its origin. We assume that our model's global configuration at time t can be reconstructed using the rules by which a unit of the system detects a pattern and scatters it out. It's a process which depends on the existence of a local set of variables that are themselves affected dynamically by the state of the unit. Preliminary findings lead us to acknowledge the constraint associated with the question of decidability. How do we know that our model can be mapped to an ICA? It was shown in a paper by Toffoli and Margolus [3] titled '*Invertible Cellular Automata: A Review*' that "There is no effective procedure for deciding whether or not any arbitrary 2-dimensional cellular Automata given in terms of a local map is invertible". The paper credits J. Kori for the proof. This theorem can be generalized to any Cellular Automata. We need to try different CA until we get to a candidate CA that can be easily mapped to an ICA. We remind ourselves that CA will be used to model pattern scattering. The difficulty lies in that while simulating patterns scattering processes described before- with CA, reconstructing the global configuration at any time from various global maps of the CA is the most important one. Again, the CA will be of great help in localizing cluster of activities among sub-units of the whole system.

Assume the following:

A pattern consisting of two sub-patterns originating from two different sources is detected at the level of the system. The pattern is encoded as a set of variables that are assumed to be highly descriptive of a known example of security compromise. These patterns propagate across a set of constructed CAs. Scattering occurs, and if the two sub-patterns do not belong to the same category a divergence occurs represented by activity in two remote areas of the corresponding CA, representing two sub-units. It is at the level of these areas that the scattered patterns stop propagating. The two areas of activity should be detected by the CA at a given time, and the corresponding map represents the relations between these two sub-patterns. It is known that

ICA are information conserving. So we may need as many global maps as configurations per unit time. That may be a huge number. Memorizing all maps and retrieving the right one is very computing-intensive. If at each time interval a map is recorded then, because the system is a continuous dynamical active one, a way to map each configuration into the right CA map must be determined.

PLANNED WORK:

In order for this modeling process to succeed, the following must be completed for each unit comprising the IS (external units that are foreign to the IS but do interact with it must be considered):

- a full survey of most human-initiated security compromises must be achieved,
- a set of corresponding and interdependent variables must be found,
- vectors whose coefficients are made up of such variables must be constructed,
- a cellular automata must be designed for each unit,
- relationships between all CA (relationship between all sub-entities) must be defined,
- the system must be programmed to map all activities in real time and ICA must be found to establish the source of a compromise when a given compromise is detected at the level of a single CA or a collection of interacting CAs.

REFERENCE

- [1] Kifer, Y., Random Perturbations of Dynamical Systems, Birkhauser, Boston, (1988).
- [2] Gutowitz, H., Cellular Automata, MIT/North-Holland, Cambridge, MA, (1991).
- [3] Tomasso Toffoli and Norman H. Margolis, "Invertible Cellular Automata: A Review," *Physica D* **45** 229-253 (1990)

Examination of the Plausibility of Network Access Compromise using USB and Live CD Tools

Lt. William Galligan
High Security Center
State of Rhode Island
Warwick, RI

Doug White
FANS Lab, CIS Networking and Security
Roger Williams University
Bristol, RI
(401) 254-3165

Abstract

Physical compromise has long been established as an avenue which leads to compromise of systems. This paper tests the plausibility of bypassing the Windows XP and Vista Operating Systems roaming profiles to access local and network drives using a machine which is rebooted containing USB flash drives and CD/DVD media. These media contain active operating systems such as Knoppix, Helix, etc. as means to compromise the system. This paper reports on the testing of such media, the outcomes of this type of compromise and develops a recommendation for the preventions such attacks in real world environments.

Paper Findings

Two approaches to examination of security compromise were utilized in this approach, the use of Knoppix Live CD and the use of the Helix bootable USB. Each of these approaches was utilized on a Windows Server to attempt to bypass the typical login security on these servers and obtain sensitive information. In both cases, it was possible to capture sensitive information as both platforms bypass the traditional login security approach used in Microsoft Servers. This is not to criticize the security level of the Microsoft servers but rather accentuates the issue of physical security of servers in general. The quintessential finding of the study was that if physical access can be obtained, it is unlikely that information can be protected short of encrypting all data on the drives.

It is of interest to determine appropriate recommendations for the securing of business servers and the determination of the ability to detect intrusions and prevent such break-ins.

Knoppix

Knoppix is a bootable, live-cd version of linux which is based on GNU/LINUX applications and kernels and contains a wide-collection of drivers to support many hardware platforms as well as a

Examination of the Plausibility of Network Access Compromise using USB and Live CD Tools -- 1

collection of systems tools which are useful in security penetration of systems. Knoppix is available from www.knoppix.org and can be burned onto CDs easily.

Helix USB

Helix linux is a Knoppix derivative which also contains a wide variety of tools and drivers to support access to many different platforms. In this case, helix 1.9 was used to create a bootable USB flash drive which was used in the compromise.

Server Compromise

Several real-world production servers were utilized in this test and in each case, the server was turned off and rebooted into the linux operating system. In all cases, this did require physical access to the servers in order to insert the CD or flash drive into the server.

Findings on Servers

The linux tools in both cases allowed the mounting of the main and secondary drives on the Windows servers with their partitions intact and viewable. In particular, it was possible through the use of the linux setup to extract database items, usernames, social security numbers, and other information which was stored on the server platforms. It was consistently possible to reveal information that would be considered sensitive and in the case of the USB flash drive it was possible to copy this information onto the flash drive for transport off site.

Recommendations

Based on the findings, it is recommended that server administrators take the following actions.

- In all cases, physical access to server areas should be restricted through the use of a combination of items. Server access should be controlled via a three prong approach recommended by CISSP, ISACA, and others (<http://www.cs.cornell.edu/Courses/cs513/2005fa/NNLauthPeople.html> for specific recommendations). This recommendation should likely include both server room access and server specific access controls.
- In all cases, server access should require BIOS setup to specifically exclude CD-ROM, USB, and Network Boot (PXE) from operation via the BIOS.
- In all cases, the BIOS should be protected via a strong password approach in the BIOS.

Examination of the Plausibility of Network Access Compromise using USB and Live CD Tools -- 2

- Physical Access to the case should be restricted. Locks on the Rack or Locks on the case should be used to prevent opening the case and physically bypassing the BIOS password through the use of battery removal or CMOS reset jumper.

A secondary recommendation is also provided:

- In the case of sensitive or privacy related data, it may be fortuitous to consider on disk encryption techniques as this will prevent viewing of the data even in the event of physical compromise (provided the encryption keys are not stored in plain text on the server). Strong encryption may be used on a partition, drive, or file to remediate this exposure. There is the risk of loss of data should the encryption be corrupted or the keys to the encryption be lost so encrypted data is a threat in and of itself and care should be taken that proper protocols are used to protect this type of data.

Evidence

The only evidence of such break-ins having occurred involved the review of the logs for system restart information. If the server could be cleanly shutdown there was a log entry for the shutdown and the reason this was done. In the event the server was hard rebooted (using a power off or reset switch) there was only evidence the server restarted in the logs. This may easily be overlooked. Investigators searching for this type of break in will be hard pressed to find evidence it occurred if systems are regularly restarted and no logs of these restarts are kept.

Conclusions

This research project provided evidence that physical compromise of server space is a constant threat and that readily available tools could be utilize to penetrate and recover sensitive data with little effort or time involved. Likewise, it was possible to leave very limited footprints of the entry having occurred and may even be possible to scrub logs in the system using the tools at hand. Thus, this research provides further support for the continued need for diligence in the physical security of servers containing any sensitive data.

Future Research

This project implies a need for further examination to determine the likelihood of such compromise and also the ability of would be attackers to scrub logs and remove all trace of this type of information espionage.

Examination of the Plausibility of Network Access Compromise using USB and Live CD Tools -- 3

THE IMPACT OF BUSINESS AND IT CHANGE ON STRATEGIC INFORMATION SYSTEMS ALIGNMENT

Henry E. Newkirk, East Carolina University, Greenville, NC, 27858-4353, (252) 328-6631, newkirkh@ecu.edu,

Albert L. Lederer, University of Kentucky, Lexington, KY, 40506-0034, (859) 257-2536
lederer@uky.edu

Alice M. Johnson, North Carolina A&T State University, Greensboro, NC, 27411, (336) 334-7656, ext. 6012, amjohns1@ncat.edu

ABSTRACT

This research investigated the effect of business and information technology (IT) change on strategic information systems planning (SISP) horizon, of horizon on SISP, and SISP on the alignment of IS strategy with business strategy. A postal survey collected data from 161 IS executives. Using structural equation modeling with PLS, business change predicted SISP horizon, but IT change did not. SISP horizon predicted SISP itself, and SISP predicted alignment. These findings provided implications for future research and practice.

INTRODUCTION

Business and IT change creates uncertainty that reduces business performance and makes the selection of new information systems all the more critical (Choe et al. 1998; Raymond et al. 1995). The process of choosing a portfolio of new IS is known as SISP (Mentzas 1997). A key temporal dimension of planning is planning horizon (Camillus 1982). It is the period of time for which the plan is developed (Das 1991). The investigation of horizon has lacked both scope and depth in strategic management research (Das 2004). Change is another critical factor in planning where research has been sparse. The potential impact of change in both business and IT demonstrates that the future can be quite incongruent with the past (Butler 1995). The purpose of this research is to test the effect of business change and IT change on SISP horizon, the effect of the horizon on SISP, and the effect of SISP on the alignment of IS strategy with business strategy. An improved understanding of the role of change, SISP horizon, and SISP may help managers better choose the new information systems that will support their organizations.

CONSTRUCTS

SISP is the process of identifying a portfolio of computer-based applications that help organizations achieve their goals (Lederer and Sethi 1988). It has been described in terms of phases and the specific tasks within them (Mentzas 1997). Business change has been defined as the rate of product/services obsolescence and the rate of product/services technology change (Miller and Friesen 1983). Change in IT refers to the differences over time in commercially available software and hardware. A planning horizon represents the longest forward planned task in the planners' active present, that is, the farthest forward being envisioned at the moment of

planning (Jaques 1982). Alignment is the linkage between business strategy and IS strategy (Baets 1992; Henderson and Venkatraman 1993; Henderson et al. 1987; King 1978). The authors used the constructs above as the basis for an instrument.

HYPOTHESES

Impact of Business Change on Strategic Information Systems Planning Horizon

Rapid business change creates uncertainty and complexity (Miller and Friesen 1980, 1982, 1983). Products and services become obsolete very quickly and technologies used to create new products and services likewise evolve quickly. Under the precautionary principle –the principle that an action potentially causing harm (such as financial waste) places the burden of proof on the advocates of change – the vulnerability would more likely inspire managers to caution rather than haste (Arrow and Fischer 1974; Epstein 1980; Gollier, Jullien and Treich 2000). New information systems can require lengthy development durations and thus longer planning horizons. Hence, we hypothesize:

H1: More rapid business change predicts longer strategic IS planning horizons.

Impact of IT Change on Strategic Information Systems Planning Horizon

Rapid IT change creates uncertainty and complexity for the organization (Benamati and Lederer 2000). The greater uncertainty and complexity associated with vendor IT change leave the end-users' organization more vulnerable, and inspire its managers to caution rather than haste wherever possible (Arrow and Fischer 1974; Epstein 1980; Gollier, Jullien and Treich 2000). Such caution would be reflected in longer SISP horizons. Hence:

H2: More rapid IT change predicts longer strategic IS planning horizons.

Impact of Strategic Information Systems Planning Horizon on the Planning Itself

“Longer planning horizons ... reflect a more insightful vision of future events” (Das 2004, 69). Such visions would press managers to devote more attention to SISP's intricate and complex group of specific, interrelated tasks which can demand much time and energy from business executives, managers, professionals and others throughout the organization. Hence:

H3: A longer strategic IS planning horizon predicts more SISP.

Because such reasoning would apply to each phase of SISP, additional hypotheses are proposed. The development of more insightful vision would encourage planners to devote more effort to determining key planning issues, defining planning objectives, organizing the planning team(s), and obtaining top management commitment. Hence:

H3a: A longer strategic IS planning horizon predicts more strategic awareness.

The development of more insightful vision would encourage planners to devote more effort to analyzing current business systems, organizational systems, and information systems, as well as to analyzing the current external business and IT environments. Hence:

H3b: A longer strategic IS planning horizon predicts more situation analysis.

The development of more insightful vision would encourage planners to devote more effort to

identifying major IT objectives, identifying opportunities for improvement, evaluating those opportunities, and identifying high level IT strategies. Hence:

H3c: A longer strategic IS planning horizon predicts more strategy conception.

The development of more insightful vision would encourage planners to devote more effort to identifying new business processes, new IT architectures, specific new projects, and priorities for new projects. Hence:

H3d: A longer strategic IS planning horizon predicts more strategy formulation.

The development of more insightful vision would encourage planners to devote more effort to defining a change management approach, defining an action plan, evaluating the action plan, and defining follow-up and control procedures. Hence:

H3e: A longer strategic IS planning horizon predicts more strategy implementation planning.

The Impact of Strategic Information Systems Planning on Alignment

SISP would reasonably be expected to produce greater knowledge about competitors, resources, customers, and regulators. Such knowledge would result in greater top management confidence and commitment. Greater commitment would result in a better plan with greater likelihood of implementation (Basu et al. 2003) resulting in greater alignment of IS strategy with business strategy. Hence:

H4: More strategic IS planning predicts greater alignment.

The effects of more meticulous planning would be expected to apply to all five SISP phases. Strategic awareness, with a more careful determination of planning issues and objectives, organizing of the planning teams, and top management commitment would, in effect, produce more alignment. Hence:

H4a: More strategic awareness predicts greater alignment.

Situation analysis, with a more careful study of the current business, organizational, and information systems, and the current external business and IT environments would provide a better foundation on which to base the plan, and enable the plan to produce more alignment. Hence:

H4b: More situation analysis predicts greater alignment.

Strategy conception, with a more meticulous identification and evaluation of opportunities, would provide more realistic alternatives. The identification of major IT objectives would enable the organization to better align future IT objectives with business objectives. The identification of high-level IT strategies would permit a basis for better final plan choices. Better alternatives and choices would enable the plan to produce better results, meaning more alignment. Hence:

H4c: More strategy conception predicts greater alignment.

Strategy formulation, with a more careful identification of the plan itself (i.e., its processes, architectures, and new projects) would provide a plan more likely to meet planning objectives. Better prioritization would result in greater likelihood of implementation and thus greater chances of meeting the planning objective of alignment. Hence:

H4d: More strategy formulation predicts greater alignment.

Strategy implementation planning, with more knowledgeable attention to change management and a better action plan would result in a greater likelihood of plan implementation. Better follow-up and control would result in a greater portion of the plan being implemented. Greater

implementation would produce better delivery of the planning objective of alignment. Hence:
H4e: More strategy implementation planning predicts greater alignment.

METHODOLOGY

This study used a field survey of IS executives. The instrument operationalized five constructs: business change, IT change, SISP horizon, SISP, and alignment. Each used items of five-point Likert scales except planning horizon, which simply asked for the number of years.

The business change construct had two items to measure the extent subjects agreed that products, services, and their technologies become obsolete or change very quickly. The items were based on Teo and King (1997) as derived from Miller and Friesen (1980, 1982, 1983) and Sabherwal and King (1992). The information technology change construct measured the extent that subjects agreed that their organization's current IT differed from its past IT and would differ from its future IT. This construct consisted of three items. They were developed and used by Benamati (1997). SISP construct measured the extent to which the organization conducted the five planning phases and their tasks. The items were derived from Mentzas (1997). The alignment construct measured the extent the organization fulfilled its IS alignment objectives. It used eight items from Segars and Grover (1998).

Pilot Test

Five IS executives were invited to participate in a pilot test. Four were Chief Information Officers, and one was Director of Information Services. Their experience ranged from 17 to 38 years. They worked in a variety of industries. Each completed the survey in about 17 minutes. After completing the instrument, each was asked to identify anything unclear or confusing. A few minor concerns about the content, length, and overall appearance of the instrument were raised. Changes from each executive were integrated into the survey before the subsequent one began filling it out. The interview with the fifth resulted in no changes.

Data Collection and Demographics

A sample of IS executives was randomly selected from a directory of top computer executives in the United States. The survey was mailed to 1,200 executives, 220 of whom returned the survey, producing an 18% response rate. Fifty-nine returned only demographic data, saying they had not taken part in an organization's SISP. The data analysis thus used the remaining 161 surveys.

Nonresponse Bias and Common Method Variance

Nonrespondents may answer survey questions differently than do respondents. This can bias survey research results. Researchers examined nonresponse bias using a time-trend extrapolation test (Armstrong and Overton 1977). The test assumes that nonrespondents resemble late respondents more than they resemble early ones. With the first 25% as early respondents and the last 25% as surrogates for nonrespondents, a multivariate analysis of variance of the 34 scaled

variables indicated no significant differences (Wilks' Lambda=.45; $p=.56$). This finding is consistent with the absence of nonresponse bias.

Common method variance is a survey research problem where individual subjects are suspected of giving socially acceptable answers. The CIO is typically viewed as the most knowledgeable person in the organization to assess SISP (Premkumar and King 1992). Most IS research has thus used a single subject to assess it (Sabherwal 1999; Lederer and Sethi 1996). The current study employed Harman's single-factor test to check for common method variance (Schriesheim 1979; Podsakoff and Organ 1986). The analysis identified ten factors with Eigenvalues greater than one, with no single factor explaining most of the variance (i.e., they ranged from 3.06% to 30.38%). These results are consistent with the absence of common method variance.

STATISTICAL ANALYSIS

Partial Least Squares (PLS) Graph version 3.0, a structural equation modeling (SEM) tool, was used to validate the measurement model and test the hypotheses (Chin 2000). PLS employs a least squares estimation procedure that places minimal demands on measurement scales, distributional assumptions, and sample size (Chin 1998; Falk et al. 1992; Fornell and Bookstein 1982; Wold 1982). Statistical significance with PLS can be assessed using a bootstrap re-sampling procedure. The current study used 500 re-samples.

Validation of the Measurement Model

The psychometric properties of the measurement model were assessed to examine internal consistency reliability (ICR), convergent validity, and discriminant validity (Chin 1998). ICR values resemble Cronbach's alpha. Values of .70 or higher are considered adequate (Fornell and Larcker 1981). Convergent and discriminant validity were assessed using two criteria. First, the square root of the average variance extracted (AVE) by a construct from its indicators. That value should be at least .707 (i.e., $AVE > .50$) and should exceed that construct's correlation with other constructs (Fornell and Larcker 1981). Second, standardized item loadings should generally be at least .707, and items should load more highly on their own constructs than on others (Carmines and Zeller 1979). After dropping two alignment items because of low factor loadings, the remaining loadings generally exceeded .707 and all exceeded their cross loadings. All ICR values exceeded .70. Each correlation between latent constructs was less than the square root of its AVE. Thus the analysis supported the reliability, and the convergent and discriminant validity of the constructs.

Hypothesis Testing

Business change and IT change were the independent variables predicting SISP horizon which in turn predicted SISP. Alignment was the final dependent variable. To test H3 and H4, SISP was a second-order construct comprising its five phases whereas to test H3a-e and H4a-e, the five phases served as first-order constructs. Results of the hypothesis testing show that H1, H3, H3a-e, H4, H4a and H4c were supported.

DISCUSSION

More rapid business change predicted longer SISP horizons (H1, $p < .05$). This finding is consistent with the expectation that rapid business change creates uncertainty and complexity, leaving managers more cautious and more inclined to use longer planning horizons.

More rapid IT change failed to predict longer SISP horizons (H2). This finding was inconsistent with the expectation that rapid IT change creates considerable uncertainty, leaving the organization more inclined to delay IT implementation decisions via longer planning horizons.

A longer planning horizon predicted more SISP for the second-order construct (H3) and for each of the specific planning phases (H3a-e). This finding supports the expectation that the development of the vision associated with longer planning horizons would produce a plan that accounts for the many possible changes during a longer horizon.

SISP predicted alignment for the second-order construct (H4) and the strategic awareness (H4a) and strategy conception (H4c) phases. The strong support for the two phases perhaps suggests the enduring effect of wisely beginning the planning efforts (strategic awareness in H4a) and the favorable effect of cleverly envisioning the alternatives (strategy conception in H4c).

Implications for Future Research and Practice

This study found support for three of its four main hypotheses, and seven of its ten SISP sub-hypotheses. Future research might investigate why the other hypotheses were not supported.

The current study used constructs of multiple-scaled items. Perhaps qualitative case research, would provide an understanding of why several hypotheses were not supported. Great interest has recently emerged concerning the impact of individual differences in the conceptualization of time (Mosakowski and Earley 2000). Some planners might be inherently long-term oriented and others might be short-term oriented. Future research might control for such differences or investigate their moderating effects on the creation of SISP horizons.

The current study suggested that business change may make managers cautious in setting SISP horizons (H1), but that IT change does not inspire such circumspection (H2). This interpretation implies that individual managers may want to assess their own degree of caution in response to business and IT change (Das 1991; Mosakowski and Earley 2000).

Support for the impact of strategic awareness (H4a) and strategy conception (H4c) on alignment reassures managers of the significance of those two phases. The failure to find the expected effects of situation analysis (H4b), strategy formulation (H4d), and strategy implementation planning (H4e) on alignment may suggest to managers that they rethink the tasks in those phases.

REFERENCES

References are available from the senior author on request.

An Examination of Factors Affecting Multilevel E-Voting

David P. Kitlan
Pennsylvania State University Harrisburg
School of Business Administration
777 West Harrisburg Pike Middletown PA 17057-4898
717-948-6639
dpk104@psu.edu

Rhoda C. Joseph
Pennsylvania State University Harrisburg
School of Business Administration
777 West Harrisburg Pike Middletown PA 17057-4898
(717) 948-6144
ruj1@psu.edu

Abstract

Electronic voting (e-voting) is a sub-component associated with the broader area of electronic government (e-government). Broadly defined, e-government refers to the use of information and communications technology (ICTs) to improve the operation of government (OECD, 2003). Specifically, e-voting refers to the use of ICTs to facilitate the voting process. The e-voting phenomenon has not been widely adopted in the United States as the standard method for voting. However, there are selected areas of interest in different levels of government. The objective of this study is to conduct an empirical analysis of the factors affecting the adoption of e-voting. The technology adoption literature forms the theoretical foundation for this analysis. We examine the impact of perceptions about government, perceptions about web technology, and individual characteristics, on the individual's intention to adopt e-voting. Our study seeks to extend the current literature on e-voting by examining e-voting at multiple levels including local, state, federal, and opinion polls instead of at a generic single level.

Introduction:

There are many issues involved in the use of information and communications technology (ICTs) to support governmental processes. These continue to dramatically change the nature of how constituents interact with government entities. Electronic government (e-government) is one of the primary ways that a government uses ICTs to have an impact on public administration and public policy issues. We begin the paper with an overview of the e-government phenomenon, and then following with an examination of e-voting. E-voting is examined from both historical and theoretical perspectives.

Next we discuss concepts related to the adoption of a technical innovation. The innovation diffusion theory (Rogers, 1995) and the technology acceptance model (Davis, 1989) are two important foundation elements in this study. Thirdly, we present a theoretical model for our study. This paper presents a research-in-progress model for the examination of the primary factors affecting the adoption of e-voting.

Overview of E-government:

In order to understand the application of technology with respect to public policy issues, it is important to consider the nature of e-government. In general, e-government can be viewed as the process of providing public value through the use of ICTs (Capati-Caruso, 2006). A recent report from the Council for Excellence in Government states that e-government “has the greatest potential to revolutionize the performance of government and revitalize our democracy” by enhancing efficiency, decreasing transaction time, bringing people closer to their government, and enhancing methods for citizens to participate in governmental affairs (Dearstyne, 2001, p. 17). Additional definitions of e-government are defined below:

1. The use of the Internet and other digital technologies to simplify or enhance the method by which citizens, employees, business partners and other government organizations interact and transact business (Koh and Prybutok, 2005).
2. The leveraging of the capabilities and power of IT to deliver services provided by governments at local, municipal, state and national levels (Grant & Chau, 2005).
3. The use of information and communication technologies, particularly the Internet, as a tool to achieve better government (OECD, 2003).
4. The use of information and communication technologies in all facets of the operations of a government organization (Koh & Prybutok, 2003).
5. The use of the Internet by governments to deliver services, to collect data, and to enhance democratic processes (Bretschneider, 2003).
6. The emerging reliance of government on digital information to make information and services available and to engage citizens in a way that meets their needs and reduces apathy and suspicion of government (Dearstyne, 2001).

Government entities and public policy administrators cannot ignore the changes that occur as a result of the use of ICTs. In the early 1990s, Freeman (1993) identified the important role that ICT would have in shaping public policy, and cautioned governments about neglecting its significance. Because e-government continues to evolve rapidly, it is important to continue to examine its adoption and impacts on different stakeholders. This study is primarily concerned with the citizen’s view of e-government. From an evolutionary perspective, e-voting is one of the later stages in the development of e-government (Siau and Long, 2005).

E-government constituents include citizens, employees, businesses, and other governments, and leads to four categories of e-government: G2C, G2E, G2B, and G2G. The first category of e-government, and the one most closely related to e-voting processes, is government-to-citizen (G2C). This refers to electronic communications and transactions that occur between a government and one or more of its citizens. Governments tend to focus on this type of interaction because a founding principle of governance is to serve its citizens. One recent international study indicates that governments around the globe recognize that a customer-centric focus is critical for e-government success (Hunter & Jupp, 2001).

Government-to-employee (G2E) initiatives involve the human resource management aspect between government and its employees. Three main benefits of G2E projects are improved strategic planning, cost reduction, and service improvements between management and employees (Ruël, et al., 2004). Government-to-business (G2B) initiatives refer to communications and transactions facilitated by electronic means between a government and a non-profit or for-profit organization. For example, the collection of corporate taxes would be a G2B process. Lastly, government-to-government (G2G) initiatives refer to entities of governments interacting with other governmental entities. G2G occurs both vertically, where information is exchanged between hierarchical levels of government, as well as horizontally, where one department interacts with another similar branch of government (Layne and Lee, 2001).

Overview of E-voting:

E-voting is often identified as one of the more advanced developmental stages of e-government. The process of e-voting can also be characterized as an occurrence of electronic democracy. Traditional democracy is considered to be the set of political processes in which citizens apply their sovereignty through their intervention in government with the goal of improving their own conditions (Moreno-Jimenez & Polasek, 2003). The future of direct democracy may be related to the number of citizens who choose to participate and how accurately their votes represent the interests of society as a whole (Morse & Hodges, 2002).

Recent forms of direct democracy are evolving into what is termed deliberative, or participative democracy. This evolution involves a transition from a passive, traditional voting system to an active, participative one. Deliberative democracy is based on a system in which citizens more actively participate to discuss a problem and accept a consensus.

Kakabadse, Kakabadse & Kouzman (2003) describe electronic democracy as an alternative model which involves “the capacity of the new communications environment to enhance the degree and quality of public participation in government” (p. 47). This work identifies four proposed models of electronic democracy: electronic bureaucracy, information management, populist, and civil society; all of which are made possible by the use of interactive information technologies.

According to Kakabadse et al. (2003), electronic bureaucracy refers to the electronic delivery of government services; the information management model of electronic democracy involves more effective communication, which connects individual citizens and decision makers; the populist model allows citizens to communicate their views on current issues; and finally the civil society model refers to the “transformation of political culture”. Kakabadse et al. (2003) analyzes each of these models in terms of their applicability and impact on democratic governance, thus offering alternative views to the area of

e-government. Conceptually, e-voting can be relevant to all four models of this electronic democracy framework.

Looking more specifically at voting processes, balloting methods have achieved significant progress since the 2000 presidential elections and changes are expected to rapidly continue (Morse, 2002). Although there have been many difficulties to overcome, e-voting has been established as a viable alternative to the varied and problematic voting methods of the past. The 2002 Help America Vote Act has provided states with some of the funding that has been needed to pursue these changes. The use of new technology has increased the accuracy, expedience, and convenience of election processes and has opened the door to additional technology upgrades in the future (Morse, 2002). There is a need to consider the future of democratic voting with respect to the use of technological progress in order to provide greater opportunities for citizen participation (Morse & Hodges, 2002).

The impact of e-voting on our democratic system is only now beginning to be researched and understood. Although there are numerous problems to overcome, involving logistics, security, secrecy, privacy, legal obstacles, equal access, and equal representation, there are also numerous potential benefits. These benefits include enhanced participation, reduced costs, and ease of registration. These potential problems and benefits are considered in more detail in Morse and Hodges (2002).

Morse and Hodges (2002) also provides a thorough review of past and present voting methods used in different countries, and they also consider the question: if political participation becomes easier and more convenient, will voting participation increase? Facing voter turnouts at or below 50% in most elections, the United States is among the lowest participating democracies of any in the world (Morse & Hodges, 2002). It may be possible to improve the level of participation by providing more voting options, as well as to make better use of technology as part of an e-voting process. Some arguments for e-voting in many states are derived from the idea that it is more convenient than traditional voting methods, and would possibly increase voter turnout (Delk, 2001).

Several states, including Arizona and California, have already taken significant steps toward achieving a phased implementation of e-voting (Morse & Hodges, 2002). Developing and implementing policies that are based on verifiable research on models related to technology diffusion and acceptance may provide a smoother transition to e-voting.

Technology adoption and governance:

The situation in which a citizen or consumer chooses an electronic service delivery method over other traditional methods is an issue involving technology adoption. Research in this area can be viewed as varying along a continuum from applying existing theories in a technology context to the development of specific technology adoption approaches (Gilbert, Balestrini & Littleboy, 2004). Gilbert and colleagues (2004) identify three main approaches that have reliable theoretical and empirical foundations for adoption. The first of these is the diffusion of innovations (DoI) approach by Rogers (1995); the second is technology acceptance models such as the technology acceptance model (TAM) by Davis (1989); and the third is the application of diffusion to technology service quality by Dabholkar (1996). A fourth construct pertaining to technology adoption and governance is that of trust. Trust, though a commonly referenced construct, can consist of several different components. The following paragraphs discuss the above four mentioned components pertaining to the adoption of a technological innovation such as e-voting.

Rogers (1995) states that, "the innovation is an idea, practice, or object that is perceived as new by an individual or other unit of adoption" (p. 11) and that adopters are classified into five groups ranging from innovators to adopters. DoI theory is often used to assess the effect of various attributes related to a particular innovation and its adoption. The concept has been applied to such fields as economics, marketing, sociology, anthropology, medicine, and others (Brown, 1981; Hagerstrand, 1967).

One important caveat about the diffusion theory is that it is not a single theory, but instead – a unified one. It is composed of a number of theories from a variety of disciplines, each focusing on a different aspect of the innovation process. In combination, these create a meta-theory of diffusion (Surry, 1997). The nature of diffusion theory includes the acceptance of new objects as well as ideas. Some consider diffusion to be the main driver of change in society (Bell, 1968). In this regard, diffusion is related to the concept of technology-as-tools as well as the concept of technology-as-organized-intelligence (Waller, 1982).

Other related models, known as substitution models, focus on the demand for technological innovations as a substitute for existing methods. One well-known and widely used diffusion model is the Bass model, which involves the timing of adoption of an innovation and first-purchase demand (Bass, 1969). Another substitution model the Fisher-Pry model (Fisher & Pry, 1971) is based on three assumptions, each of which is applicable to e-voting adoption. The three assumptions are as follows: 1. many technological advances can be considered as viable substitutions for other products; 2. often new technologies completely replace previous ones; and 3. the rate of substitution of new technologies for older ones is proportional to the amount of the old technology that remains.

Having considered the diffusion of innovation approach, the second main approach to technology adoption is TAM, which is an extension of the theory of reasoned action (TRA) to technology. The TRA is from the social psychology literature (Ajzen & Fishbein, 1980) and involves an individual's evaluation of the potential to perform a specific task. The TAM model proposes that beliefs have an impact on attitudes about information technology, which then lead to intentions and behaviors related to actual technology usage. The beliefs that predict the use of technological systems include perceptions about the usefulness of the technology related to improved performance, and perceptions about the ease of use of the technology (Davis, 1989).

Although most research related to technology adoption and the use of the Internet considers the positive effects of other factors on this behavior, O'Cass and Fenech (2003) have studied the factors that discourage individuals from adopting the technology. There is an increasing recognition of the need to research attitudes with respect to Internet-related adoption. Finally, several of the attitude-based theories, such as the TRA, theory of planned behavior, and theory of trying, have been integrated with external factors, such as perceived risks, to try to explain why individuals may prefer self-service options that are based on the use of technology (Bobbitt & Dabholkar, 2001). Each of these theories has some relevance, either direct or indirect, to the adoption of e-voting systems.

In the first two approaches, DoI and TAM, the perceptions of potential users and adopters of technology are beliefs that determine the behavior about a product, service or technology. Their focus is on perceptions of technology and their effect on the intent to use or adopt that technology. More recently, a third approach has been developed that involves using intentions of individuals based on service quality to explain service delivery through

the use of technology. With service quality, perceptions can relate to the evaluation of service performance after the technology is used.

In a study of consumer evaluation of self-service delivery through technology, Dabholkar (1996) proposes two models to determine the impact of service quality on the intention to use the technology. One model is based on various attributes of quality while the other is based on predetermined attitudes about the technology. The attribute model uses dimensions that are similar to those used in other service quality literature. The results of this study show that speed of delivery, ease of use, reliability, enjoyment, and control were all significant factors in assessing expected service quality. Other comparative models show that consumers compare the innovative technology service delivery with traditional alternatives (Meuter et al., 2000; Szymanski & Hyse, 2000).

The fourth approach to technology adoption involves trust theory. Trust can be defined along two dimensions: (a) as an assessment of a current situation, or (b) as an innate personality trait or predisposition (Driscoll, 1978). Trust is an important aspect of user decision making. For example, one's level of trust is an important factor affecting purchase or transaction decisions. More recently, trust has been examined with respect to electronic commerce (e-commerce) (Jarvenpaa, Tractinsky, & Vitale, 2000; Koufaris & Hampton-Sosa, 2004). A citizen that has previously not established trust in the area of e-commerce may transfer that lack of trust to other electronic domains such as e-government.

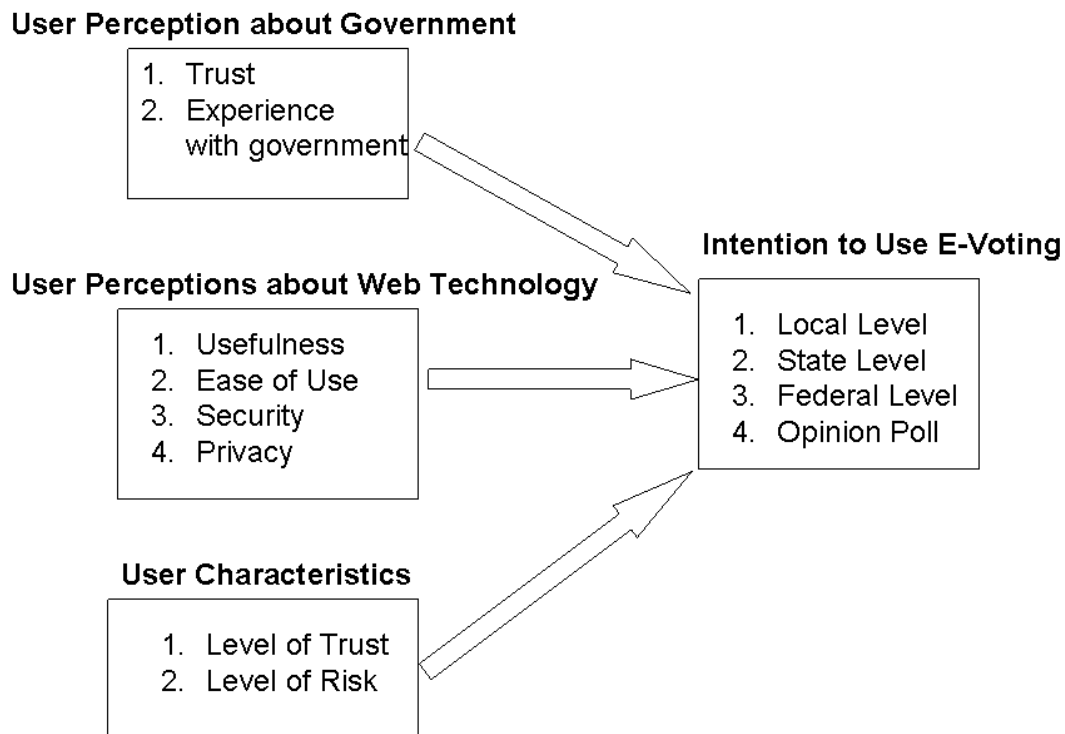
The successful diffusion and acceptance of public administration functions via e-government requires two levels of trust: trust of the government and trust of the technology that supports the interaction (Carter & Belanger, 2005; Lee & Turban, 2001). Thus, if a constituent has limited trust in either the technology or government, the use of an e-government process such as e-voting, can be limited. Trust, along with financial security, are two critical factors limiting the adoption of e-government services (Gilbert et al., 2004). Therefore, it is important to maintain effective security within e-government systems to promote and protect consumer trust and confidence.

Researchers have recently begun to integrate the previous approaches into models to identify the major factors influencing the adoption of online government services by citizens on various levels. The integration of approaches can reduce the limitations of the individual approaches (Gilbert et al., 2004). Carter and Belanger (2005) have developed a comprehensive e-government adoption model that combines constructs from DoI theory (Rogers, 1995), TAM (Davis, 1989), and web trust theory (McKnight et al., 2002). Schaupp and Carter (2005) have extended this framework to explore the intention of citizens to use an online voting system. The results indicate that user perceptions of compatibility, usefulness, and trust significantly impact the intention of young citizens to use an e-voting system (Schaupp and Carter, 2005).

Theoretical Framework and Proposed Methodology:

We propose a theoretical framework (figure 1) for the examination of an individual's intention to use e-voting. Voting, and ultimately e-voting can occur at several levels in the United States, including local, state, and federal. As part of this study, we propose that voting at each of these three levels has varying degrees of significance to the citizen. Voter turnout in local level elections is generally lower than that in federal level elections. Additionally, through the use of ICT, citizens can participate in the democratic process by participating in online polls that are not necessarily associated with electing an official. Based on this information a citizen's intention to use e-voting can be influenced by the level and/or impact of the election.

Figure 1: Multilevel E-voting Framework



There are three main factors affecting the user's intention to use e-voting. The first factor is user perception about government. Perception about government has two components: trust and experience. We use experience to determine whether the individual has had an overall positive or negative experience with government. The second variable in the model is perception about web technology. The four relevant components are: usefulness, ease of use, security, and privacy. Security and privacy are two factors that recur as concerns for individuals using web technologies, and they are particularly relevant to voting.

Lastly, we expect that individual user characteristics will influence their decision to use e-voting. The two characteristics included in our framework are level of trust and risk propensity. A person's innate sense of trust will also affect their trust of others. Also, a risk-averse individual may be less likely to try an innovation, such as e-voting.

The development of a valid instrument for this proposal has been adapted from previous related work including but not limited to measures of PU and PEOU from Davis (1989); intention to use adapted from Van Slyke et al. (2004) and Davis (1989); measures of trust in government and trust in technology have been adapted from Van Slyke et al. (2004) and McKnight et al. (2002). Demographic information such as age, income, education, and technology experience will also be collected as part of the user background information.

We will pilot the instrument on a sample group of 25 users to determine if it is clear and understandable. Once we have resolved any associated design limitations we plan to distribute our survey instrument to a population of approximately 200 undergraduate students. A quantitative methodology will be used in the collection and analysis of the data.

Limitations and Conclusion:

E-voting will likely receive more attention, both in research and in practice, as the supporting technologies are developed. The objective of this study is to examine the factors affecting the adoption of e-voting at multiple levels. However, there are several limits of our proposed study. Our first challenge is the use of a student population for the data collection since college students may not represent the traditional voting demographic. However, to overcome this limitation we suggest that students are very technology literate, and may ultimately be some of the early adopters of e-voting systems. Secondly, our study is a survey with no direct incentive for user participation. Ultimately, there may be some selection bias by the respondents that fill out the survey.

E-voting is an area of governance that will likely have far reaching implications in the future. In spite of the identified limitations, this study can provide important insights to both researchers and practitioners in the field. E-voting may soon become one of the main ways to link electronic government and the implementation of public policy, and it will continue to attract the attention of researchers, policy analysts and other stakeholders in the public sector in the future. It is hoped that this proposed research, and related efforts, will produce the information necessary to support a successful transition to electronic governance and the use of e-voting.

References

- Ajzen, I. & Fishbein, M. (1980). *Understanding Attitudes and Predicting Social Behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Bass, F. M. (1969). A new-product growth model for consumer durables. *Management Science*, 15(Jan), 215.
- Bell, D. (1968). *Toward the year 2000: Work in progress*. Boston: Beacon Press.
- Bobbitt, L.M. & Dabholkar, P.A. (2001), Integrating attitudinal theories to understand and predict use of technology-based self-service. *International Journal of Service Industry Management*, 12(5), 423-50.
- Bretschneider, S. (2003). Information technology, e-government, and institutional change. *Public Administration Review*, Nov/Dec 2003, 63(6), 738.
- Brown, L. (1981). *Innovation diffusion: A new perspective*. London: Methven.
- Carter, L. & Belanger, F. (2005). The utilization of e-government services: Citizen trust innovation and acceptance factors. *Information Systems Journal*, 15(1), 5-25.
- Dabholkar, P.A. (1996), Consumer evaluations of new technology-based self-service options: An investigation of alternative models of service quality. *International Journal of Research in Marketing*, 13(1), 29.
- Davis, F. (1989), Perceived usefulness, perceived ease of use and user acceptance of information technology. *MIS Quarterly*, 13(3), 319.
- Dearstyne, B. W. (2001). E-business, e-government and information proficiency. *Information Management Journal*, Oct. 2001, 35(4), 16.
- Delk, K. C. (2001). What will it take to produce greater American voter participation? Does anyone really know? *J. Pub. Int. L.*, 133, 175.
- Driscoll, J. W. (1978). Trust and participation in organizational decision making as predictors of satisfaction. *Academy of Management Journal*, 21(1), 44-56.
- Fisher, J. C. & Pry, R. H. (1971). A simple substitution model of technological change. *Technological Forecasting and Social Change*, 3(Mar), 75-88.
- Freeman, C. (1993). Technical change and future trends in the world economy. *Futures*, 25(6), 621-635.
- Gilbert, D., Balestrini, P. & Littleboy, D. (2004). Barriers and benefits in the adoption of e-government. *The International Journal of Public Sector Management*, 17(4/5), 286.
- Grant, G. & Chau D. (2005). Developing a generic framework for e-government. *Journal of Global Information Management*, Jan. 2005, 13(1), 1.
- Hagerstrand, T. (1967). *Innovation Diffusion as a Spatial Process*. Chicago: University of Chicago Press.
- Hunter, D. R., & Jupp, V. (2001). *Rhetoric vs reality - closing the gap*: Accenture.
- Jarvenpaa, S. L., Tractinsky, N., & Vitale, M. (2000). Consumer trust in an Internet store. *Information Technology and Management*, 1(1-2), 45.
- Kakabadse, A., Kakabadse, N. K. & Kouzman, A. (2003). Reinventing the democratic governance project through information technology? A growing agenda for debate. *Public Administration Review*, January/February 2003, 63(1), 44.
- Koh, C. E. & Prybutok, V. R. (2003). The three-ring model and development of an instrument for measuring dimensions of e-government functions. *Journal of Computer Information Systems*, 2003, 33(3), 34.
- Koufaris, M., & Hampton-Sosa, W. (2004). The development of initial trust in an online company by new customers. *Information & Management*, 41(3), 377.

- Layne, K., & Lee, J. (2001). Developing fully functional e-government: A four stage model. *Government Information Quarterly*, 18(2), 122-136.
- Lee, M. K. O. & Turban, E. (2001). A trust model for Internet shopping. *International Journal of Electronic Commerce*, 6(1), 75-91.
- McKnight, H., Choudhury, V. & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research*, 13(3).
- Meuter, M. L., Ostrom, A. L., Roundtree, R.I. & Bitner, M. J. (2000). Self-service technologies: Understanding customer satisfaction with technology-based service encounters. *Journal of Marketing*, 64(3), 50-64.
- Moreno-Jimenez, J. M. & Polasek W. (2003). E-democracy and knowledge: A multicriteria framework for the new democratic era. *Journal of Multi-Criteria Decision Analysis*, 12, 163-176.
- Morse, R. (2002). Electronic voting: Progress over setbacks. *Law Technology*, 4th Quarter 2002, 35(4), 6.
- Morse, R. & Hodges, M. (2002). E-voting and democracy: Past, present and future. Is e-voting a possibility? *Law Technology*, 3rd Quarter 2002, 35(3), 31.
- OECD. (2003). The case for e-government: Excerpts from the OECD report: "The e-government imperative". *OECD Journal on Budgeting*, 2003, 3(1), 62.
- O'Cass, A. & Fenech, T. (2003). Web retailing adoption: Exploring the nature of Internet users' web retailing behaviour. *Journal of Retailing and Consumer Services*, 10(2), 81-94.
- Rogers, E. M. (1995). *Diffusion of innovations* (4th ed.). New York: The Free Press.
- Ruël, H., Bondarouk, T., & Looise, J. K. (2004). E-HRM: Innovation or irritation. An explorative empirical study in five large companies on web-based HRM. *Management Revue*, 15(3), 364-380.
- Schaupp, L. C., & Carter, L. (2005). E-voting: From apathy to adoption. *Journal of Enterprise Information Management*, 18(5-6), 586-602.
- Siau, K. and Long, Y. (2005). Synthesizing E-government Stage Models: A Meta-synthesis based on Meta-ethnography Approach. *Industrial Management & Data Systems*, 105(3/4), 443-458.
- Surry, D. W. (1997, February). *Diffusion theory and instructional technology*. Paper presented at the Annual Conference of the Association for Educational Communications and Technology (AECT), Albuquerque, New Mexico.
- Szymanski, D. M. & Hyse, R. T. (2000). E-satisfaction: An initial examination. *Journal of Retailing*, 76(3), 309-22.
- Van Slyke, C., Belanger, F., & Comunale, C. L. (2004). Factors influencing the adoption of web-based shopping: The impact of trust. *Database for Advances in Information Systems*, 35(2), 32-50.
- Waller, W. (1982). The evolution of the veblenian dichotomy: Veblen, Hamilton, Ayres, and Foster. *Journal of Economic Issues*, 16(Sep), 757-771.

CAN YOU REALLY SHARE WHAT YOU DON'T KNOW YOU KNOW? ON THE FUTILITY OF CONTROLLING KNOWLEDGE SHARING IN ORGANIZATIONS

José Castillo

The University of Texas of the Permian Basin, School of Business, 4901 E. University Blvd.,
Odessa, TX, 79762, castillo_j@utpb.edu, (432)366-4468

ABSTRACT

Lin's argument that sharing tacit knowledge is of the utmost importance to organizations in the "knowledge economy" is novel for its subject matter and the constructs at the core of his argument. However, if we consider that at the heart of such a view is the notion of control of information coming from or given to some central machine, normally called the knowledge management system, paradoxically, we end up with some real conundrums. This paper details these issues.

Keywords: Tacit knowledge, Knowledge economy, Ethical behavior

Lin's [14] piece in the *Journal of Business Ethics* that proposed and tested for a new ethic of sharing in the "knowledge economy," is novel for its subject matter and the constructs at the core of his argument. Professor Lin's attempt to begin such an ethics discussion is seminal and praiseworthy, given "sharing knowledge" is today of interest to both, academics and practitioners in the realm of knowledge management, among others. That said however, Lin's paper suffers serious shortcomings in its constructs, argumentation, and its methodology, that make his conclusions suspect at best, if not altogether wrong. This paper sheds light on these fatal flaws.

Lin's Arguments and Conclusions

The core of Professor Lin's argument echoes the central tenet in the knowledge management literature; namely, that the viability of organizations in the knowledge economy hangs heavily on the knowledge the firm's employees possess and which they most ably share and use [1] [2] [9] [10] [11] [17] [22] [26]. And while process and product innovation are derived directly from *applied* knowledge, *managed* knowledge presupposes *shared* knowledge; thus it is, that sharing becomes the pivotal activity in the organization. Working on this assertion, Lin proposes a model and empirically verifies that the sharing of a specific type of knowledge, tacit knowledge, is mediated by two important behavioral variables; organizational commitment and trust in co-workers. Given his findings, Professor Lin concludes that "...low tacit knowledge sharing is likely attributed to a lack of organizational commitment and trust in co-workers which may be impelled by negative organization-person influence, personal influence, and interpersonal influence" [14, p.421]. And he further asserts that "...employees who refuse to practice such sharing...are perceived as latently unethical" [14, p.421].

Mixed Constructs

Critical to Lin's argument is the distinction between two forms of knowledge: explicit knowledge and tacit knowledge. But beyond the confused and unhelpful characterizations of explicit knowledge as "general information and knowledge" [14, p.413] and tacit knowledge as "job tips or influential knowledge" [14, p.413], the reader is never given a clearer division or

definition of either of these two forms of knowing. What's more, Professor Lin's sketchy characterizations belie and essentially dismiss a legacy of literature in management and philosophy that questions the idea and its understanding (the seminal piece being Turner [23]). Indeed, Polanyi's [19] original characterization of human knowing in his now famous assertion that tacit knowledge means "...we can know more than tell" [19, p.4] is particularly relevant to what Lin is attempting to argue. That is, by this definition it is quite evident that the transmission of knowledge from person to person, that Professor Lin argues should be the reigning ethic in the organization, is inherently incomplete; for as Bordum [3] puts it, "...even though I may be able to put some of my thoughts and reflections into words, there are always some things we cannot explicitly tell of on which our knowledge relies" [3, p.50].

And though Polanyi [19] never actually proposed such a thing as explicit knowledge, by proposing the very existence of tacit knowledge he implicitly gave rise to it; for inherently, to reduce knowledge to its ineffable kind means there must exist its opposite, knowledge that is easily "describable," effable. And ever since then, anyone and everyone distinguishing a tacit form knowledge is served best by explaining that, by contrast, there exists "...knowledge that can be spelled out or formalized..." [6, p.381] or "...codified and easily translated [into] facts and information" [21, p.347]. In understanding this distinction between tacit and explicit knowledge, there arises the problem for Lin and all others attempting to "manage" knowledge; for anyone verbalizing what they know, how does the speaker (never mind the listener!) know that he or she is divulging what they know tacitly? From another perspective, the problem for the organization essentially boils down to the fact that "...despite querying, testing, and surveying even the most helpful and divulging employees, any effort to harness an employee's stock of tacitly held knowledge is pointless, as this knowledge cannot be differentiated from the explicit knowledge that that person holds" [4, pp.47-48]. Arguably, Lin's "knowledge sharing" organizations in the new economy must be forced to conclude that all knowledge they possess must be of the explicit kind—this by virtue of the fact that it is strictly explicit knowledge the only knowledge we are conscious of, and therefore, that which we are most able to articulate. By extension, it would seem that anything shared between firm employees is simply information and not knowledge, and Pfeffer and Veiga [18] seem to argue this when they state that, "...if expert knowledge has a substantial component of tacit knowledge, it will be impossible for experts to present the real basis of their judgments and decisions. Experts are more likely to rely on those factors and evidence that are available and accessible to all. In doing so, forced to explain decisions to a wider audience, the experts will have to rely on the same data and decision processes as anyone else." [18, p.45]. The confusion of terminology in Lin's study is further exacerbated by his muddled view of knowledge and information.

A fundamental concept in management information systems is that 'information' is *simply* 'processed data'; that is, raw facts are summarized, tallied, sorted or manipulated in some way so as to create meaning to a reader. By extending this argument, it is conceivable that 'knowledge' is *complexly* 'processed information;' *complexly* in the sense that the processing takes place in a largely unknown machinery, namely, our minds. Supporting this assertion, Roberts [20] states that, "...knowledge is distinct from information: indeed, it is more than information, since it involves an awareness or understanding gained through experience, familiarity or learning...[moreover, it] involves cognitive structures that can assimilate information put into wider context, allowing actions to be undertaken." [20, pp.100-101]. And this point is echoed by Kane [12], who sees "...knowledge as the application and productive use of information" [12, p.34].

With knowledge and information framed this way, the terms knowledge and information cannot be used interchangeably as Lin tends to do. That is, any description of knowledge as information must contain a modifier, one expressing *processing* of some kind. Yet, Lin describes *shared* knowledge as "...individuals sharing relevant experiences and information with other organizational members...[and] despite the increasing ease of use of having access to online information, employees still must count on their co-workers for...[it]" [14, p.411]. Clearly, while knowledge may involve experience (as in learning from failure), as we've seen knowledge is much more than that; it is procedural, it is contextual, and it is applicable, thus, the focus of Lin's ideation in this case is *not* knowledge but *rather* information. Further evidence of this is found in Lin's instrument for assessing tacit knowledge. Suspending for a moment judgment as to whether four simple queries actually reveal knowledge, the subject of these questions are the study subjects' job experience, expertise, ideas, and tips, again, such naiveté about the nature of tacit knowledge belies extensive literature that argues that the transfer of tacit knowledge (if such a thing should truly be possible) occurs only through "learning by doing" as explained in another seminal piece Lin fails to acknowledge, Nonaka and Takeuchi's *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation* [16]. Lastly, in reading the questions in Lin's instrument and the confused terminology, one is led to conclude that the fundamental problem in his study lies in the operant definition that seems to pervade it; that tacit knowledge is merely universalized knowledge, that it is knowledge that has not been brought to the fore and that it is a matter of prompting to get to it. Here again, by the foregoing arguments this could not be further from the truth.

Problematic Elements of Lin's Methodology

Beyond problems with constructs and fundamental arguments, one finds certain troublesome elements in Professor Lin's study methods. For instance, while one would agree that using "experienced" subjects would help to generalize Lin's findings to a practical setting, Lin's problem is not external validity but construct validity. Moreover, experience notwithstanding, the very fact that students are the subject of Lin's study presumes inexperience in some realms (otherwise they would not be college students). And this would imply incompleteness of the tacit knowledge Lin is trying to assess. Perhaps a sounder case would be made had Lin availed himself not of students in his vicinity, but faculty. Indeed, the case for the assessment of tacit knowledge could have been strengthened by not only using different subjects but also by a more detailed and rigorous instrument such as Wagner's as detailed in Wagner and Sternberg's [25] study of industry managers, that fundamentally argues that only seasoned managers are possessed of tacit knowledge.

Another problem lies in the model tested using SEM (structural equation modeling). In his model, Lin proposed cooperativeness and instrumental and expressive ties as indicators of tacit knowledge sharing, which are mediated by organizational commitment and trust in co-workers. In support of this view, Lin states "...social network ties thus formed after individuals are acquainted with their co-workers facilitate knowledge sharing by providing for mutual support, the exchange of information, common ground for establishing shared expectations" [14, p. 417]. But by arguing cooperativeness as a separate indicator in his model, Lin implicitly seems to argue that the necessary "acquaintance between individuals" on the job takes place separate from a friendly and cooperative environment. It is perhaps more realistic to consider cooperativeness a precursor to network ties, as an indicator of the instrumental and expressive ties that are established in the workplace.

Conclusion

When it comes to human experience, language is limited in describing the variety, depth, and complexity of what builds in our minds. Indeed, as old as human expression must be, the best civilization has been able to do is to devise language that speaks to its own limits; terms like *indescribable*, *ineffable*, or sayings like *words can't describe*, and *a picture tells a thousand words* are characteristic. In view of this inescapable nature, perhaps Lin's effort and all other efforts that originate with the premise that knowledge can be managed are fundamentally flawed. And thus, rather than expending money and energy in the unattainable quest to obtain that which can only reside in the employees' minds and only they can use, firm managers might be better served by accepting the fact "...that tacit knowledge is an individual's private good. It is her own to exploit or disclose, *no matter how, when, where or from whom she has acquired or developed it*" [24, p.287]. Put another way, "...[sharing] knowledge work is...[absolutely the employees]...discretionary behavior in organizations" and so it might be more appropriate for management to focus on "...organizational characteristics such as transformational leadership, job design, social interaction and organizational culture" that create the right settings to get the best from their people [13, p.287]. This sentiment is echoed by Pfeffer and Veiga's [18] broader view of what it takes to be successful in the new economy, they argue that:

"...the key to managing people in ways that lead to profits, productivity, innovation, and real organizational learning ultimately lies in the manager's perspective...do they see intelligent, motivated, trustworthy individuals-the most critical and valuable strategic assets their organizations can have?...Do they see...[employees]...as the fundamental resources on which their success rests and the primary means of differentiating themselves from the competition? With the right perspective, anything is possible. With the wrong one, change efforts and new programs become gimmicks, and no army of consultants, seminars, and slogans will help" [18, pp.47-48].

Pfeffer and Veiga [18] further argue that perhaps the failure has been not in employees who share or do not share, but in organizations that, themselves do not share. For it is organizations that make the effort to be transparent, to share their "financial performance, strategy, and operational measures," not to mention sharing the wealth thru top compensation packages, that either directly or implicitly tell their employees that they are trusted, that they belong, that the operant approach is freedom and minimal managerial obtrusiveness. Only then can organizations operate at peak performance and generate the results they seek.

In conclusion, history refutes Lin's sharing ethic and Ebejer and Morden's [7] *limited paternalism* (i.e., the limited "...infracton on liberty...or autonomy" by one transacting party on another, lesser knowing or unaware party, [7, p.337]), for this country was built on knowledge in disequilibrium. Our nation attained its superiority by unshared *know-what*, where parties acted for their own ends but where the Unites States benefited perhaps from far more accurate knowledge (irrespective of the treatment of the previous owners afterwards). Manhattan was purchased for the equivalent of \$24 dollars in trinkets from the unsuspecting native Americans who resided there; Jefferson more than doubled the size of the United States by essentially buying the middle of America from the French for about three cents an acre in 1803; and the riches of Alaska would not be possible today if not for President Andrew Johnson's willingness to pay the Russians all of about two cents per acre in 1867. Thus it may be that the ancient latin phrase, *caveat emptor* or "let the buyer beware" is appropriate and the operant philosophy

today, where the means to uncovering any and all relevant information to a transaction is available to anyone, and is only limited by the ingenuity of the individuals seeking it.

References

- [1] Akhavan, P., Jafari, M., and Fathian, M. (2006). Critical success factors of knowledge management systems: a multi-case analysis. *European Business Review*, 18(2): 97-113.
- [2] Alstete, J. W. (2006). Knowledge management and higher education: a critical analysis. *Knowledge Management Research & Practice*, 4(3): 252-253.
- [3] Bordum, A. (2002). From tacit knowing to tacit knowledge—emancipation or ideology? *The Critical Quarterly*, 44(3): 50-54.
- [4] Castillo, J. (2002). “A Note on the Concept of Tacit Knowledge.” *The Journal of Management Inquiry*, Vol. 11(1): 46-57.
- [5] Collins, C. J. and Smith, K. G. (2006). Knowledge exchange and combination: The role of human resource practices in the performance of high-technology firms. *Academy of Management Journal*, 49(3): 544-560.
- [6] Cook, S. D. N. and Brown, J. S. (1999). Bridging epistemologies: The generative dance between organizational knowledge and organizational knowing. *Organization Science*, 10(4): 381-400.
- [7] Ebejer, J. M. and Morden, M. J. (1988). Paternalism in the marketplace: Should a salesman be his buyer’s keeper? *Journal of Business Ethics*, 7(5): 337-339.
- [8] Inkpen, A. C. and Tsang, W. K. (2005). Social capital, networks, and knowledge transfer. *Academy of Management Review*, 30(1): 145-165.
- [9] Hariharan, A. (2005). Critical success factors for knowledge management. *Knowledge Management Review*, 8(2): 16-19.
- [10] Hendriks, P. H. J. (2005). Knowledge management in organizations: a critical introduction. *Management Learning*, 36(4): 535-539.
- [11] Jackson, M. C. (2005). Reflections on knowledge management from a critical systems perspective. *Knowledge Management Research & Practice*, 3(4): 187-196.
- [12] Kane, H. C. M. (2003). Reframing the knowledge debate, with a little help from the Greeks. *Electronic Journal of Knowledge Management*, 1(1): 33-38.
- [13] Kelloway, E. K. and Barling, J. (2000). Knowledge work as organizational behavior. *International Journal of Management Reviews*, 2(3): 287–304.

- [14] Lin, C.-P. (2007). To share or not share: Modeling tacit knowledge sharing, its mediators and antecedents. *Journal of Business Ethics*, 70: 411-428.
- [15] Miller, K. D., Meng, Z., and Calantone, R. J. (2006). Adding interpersonal learning and tacit knowledge to March's exploration-exploitation model. *Academy of Management Journal*, 49(4): 709-722.
- [16] Nonaka, I. and Takeuchi, H. (1995). *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. New York: Oxford University Press, USA.
- [17] Park, H., Ribiere, V., and Schulte Jr., W. D. (2004). Critical attributes of organizational culture that promote knowledge management technology implementation success. *Journal of Knowledge Management*, 8(3): 106-117.
- [18] Pfeffer, J. and Veiga, J. F. (2001). Putting people first for organizational success. *The Academy of Management Executive*, 13(2): 37-48.
- [19] Polanyi, M. (1966). *The Tacit Dimension*. Garden City, NJ: Doubleday Anchor.
- [20] Roberts, J. (2001). The drive to codify: Implications for the knowledge-based economy. *Prometheus*, 19(2): 100-116.
- [21] Smith, K. G., Collins, C. J. and Clark, K. D. 2005. Existing knowledge, knowledge creation capability, and the rate of new product introduction in high-technology firms. *Academy of Management Journal*, 48(2): 346-357.
- [22] Tomassini, M. (2004). Understanding knowledge management: critical and postmodern perspectives. *Management Learning*, 35(3): 356-359.
- [23] Turner, S. P. (1994). *The Social Theory of Practices: Tradition, Tacit Knowledge and Presuppositions*. Chicago: University of Chicago Press.
- [24] Van Caenegem, W. (2005). Inter-firm migration of tacit knowledge: Law and policy. *Prometheus*, 23(3): 285-306.
- [25] Wagner, R. K. and Sternberg, R. J. (1985). Practical knowledge in real-world pursuits: The role of tacit knowledge. *Journal of Personality & Social Psychology*, 49(2): 436-458.
- [26] Wong, K. Y. (2005). Critical success factors for implementing knowledge management in small and medium enterprises. *Industrial Management + Data Systems*, 105(3/4): 261-279.

Critical Research in Information Systems: Making a Nest

Anil Gurung
Neumann College
Division of Business & Information Management
One Neumann Drive
Aston, Pennsylvania 19014
Email: agurung@acm.org
Telephone: (610)558-5535

Anurag Jain
Salem State College
Bertolon School of Business
352 Lafayette Street
Salem, MA 01970
Email: ajain@salemstate.edu
Telephone: (978)542-6894

Abstract

Although the dominant epistemological paradigm in IS research has been positivism, other approaches have also been espoused by researchers to explain various phenomena in Information Systems. The Critical view, in addition to Interpretive view, provides an additional lens to study the delicate intricacies of IS phenomenon. In the critical research approach, the researchers are not mere observers but they also believe that they can influence or can be influenced by the social and technological systems which are the subject of their inquiry. This paper examines the appropriateness of critical perspective in IS research and reviews the areas that have been studied in IS research using this view. Additionally, potential areas that are suitable to be understood from a Critical perspective have been discussed.

Introduction

The positivist view in social science research including Information Systems has faced many criticisms. Nevertheless, it still remains, the most dominant research perspective. Beginning in early 1990s research in IS has seen few instances with different perspectives than positivist. Interpretive is another major approach seen in IS researches. One major advantage of Interpretive over positivism since it recognizes the “life-world.” The field of phenomenology defines the “life-world” as the world of consciousness and humanly created meanings (Ngwenyama and Lee 1997). In the critical research approach, the researchers are not mere observers. They believe that they can influence or can be influenced by the social and technological systems, which are the subject of their inquiry. Critical researchers, beyond explanations and understandings of situations, are expected to critique the unjust and inequitable conditions of the situation from which the people need to be emancipated ((Ngwenyama and Lee 1997).

Information Systems (IS), being a young discipline, has seen much discussions about theoretical foundations of the various IS research streams. Hirschheim and Klein (1989) have classified the general theoretical assumptions of different approaches to IS development research into following four paradigms: Functionalism, Social Relativism, Radical Structuralism, and Neohumanism. Among these paradigms, Neohumanist paradigm has only seen limited attention in IS research. One of the approaches of this paradigm is the critical social theory. As pointed out by Ngwenyama (1991), the potential contribution of critical social theory to new knowledge in IS development has been discussed by many researchers (Mingers 1981, Klein and Lyytinen 1985, Orlikowski and Baroudi 1991, Ngwenyama 1991, Hirschheim and Klein 1994, Ngwenyama and Lee 1997).

Research issues in IS are quite diverse since it is a multidisciplinary field that includes information technology, management science, organization science, psychology, engineering, and economics. One research approach would not be comprehensive enough to solve all the issues in IS research. Most of the leading researchers have called for pluralistic approach in IS research. In that respect, critical perspective provides an additional lens to increase the understanding of problems faced in IS.

This report starts with a discussion of epistemology and ontology of various IS research approaches. Then it will give an insight into historical perspective of critical

theory, and goes on to literature review of Critical research philosophy. Afterwards, we will discuss the main tenets of this perspective followed by its critical evaluation.

Philosophical discussion of IS research approaches

In this section, the three main research approaches in IS research is examined with respect to their ontological and epistemological beliefs.

The positivist approach in IS involves the use of the rules of formal logic and the rules of hypothetico-deductive logic to inductively draw inferences about observations with regards to theoretical propositions. These theoretical propositions need to satisfy the four criteria of falsifiability, logical consistency, relative explanatory power, and survival. Ontologically, the positivist views an objective physical and social world that exist independent of human perception. The goal is to understand IS phenomenon by effectively using the right set of constructs and instruments. The assumption about social reality is that human action is assumed to be intentional and rational. Positivist belief that scientific inquiry is value-free so there is no room for subjectivity. In IS research, this belief is taken in a less constrictive view because of close relation between theory and practice. The main goal of IS research is to inform and improve the development and use of IS in organizations (Orlikowski and Baroudi 1991).

Interpretive approach states that social world has intersubjectively created meanings and therefore, need to be studied in a different way than the physical world. They interpret their empirical results in terms of what it means to the observed world (Lee 1991). Ontologically, interpretive maintains that the social reality is produced and reinforced by humans through their interactions. The epistemological belief of this approach is that social process cannot be explained in terms of hypothetical deductions, and therefore requires researchers to get involved in the social world of the phenomenon of inquiry. Interpretations about reality can also undergo changes as new meanings are formed and transferred. So, there may be multiple realities. Since the main tenets of reality is different from the one assumed by positivist, knowledge is obtained from different methodologies used in hermeneutics, ethnography etc.

Ontologically, critical approach believes that social reality is historically created and conditioned (Jonsson 1991). All the social structures including humans, organizations and relations are not limited to a certain structure. They undergo changes. Each entity has an unfulfilled potentiality, if made aware thereof, can act to influence their social structures and improve ways to realize such potential. Critical researchers take a holistic approach to observing entities and phenomenon. An entity exists only in terms of its relation to entirety and never by itself. Such a dialectic relation is historically constituted and may initiate changes. The epistemological belief of this perspective is that knowledge is based on social and historical practices. Therefore, the research with critical approach tends to be longitudinal. Critical researcher takes on a more active role than others since their goal is to effect changes in the phenomenon of study besides understanding them.

Historical perspective of Critical theory

Critical theory has its origins in Post-Marxist critical studies of society that started in the middle years of 20th century. The thoughts of Horkheimer, Adorno, Marcuse and Habermas have greatly impacted the development of critical theory. Their work draws upon some aspects of work from Kant and Hegel, but not all the aspects. They reject Kant's transcendental method and many aspects of Hegel's philosophy (Held 1980).

Unlike Marxian concern with the issues of Capitalism and calling for revolution, the critical theorists focus more on the issues of capitalist institutions that produce "alienating" forms of life. Amongst critical theorists, the works of Habermas has seemed to gain a wider recognition. He seeks to offer a theory that will emancipate humans from their social problems. Critical theory attempts to provide a source that will construct a better world in comparison to positivistic science that manifests false knowledge claims. Habermasian concept of communicative action has been given extensive consideration in critical social studies. In this concept, "reason" which is observed in terms of intersubjectivity of mutual understanding and recognition, provides a basis for seeking knowledge. His reconceptualization of action and rationalization observes rational action as a source of liberating force rather than oppressive.

The main of contention of Habermas like other critical theorists is that knowledge is historically based and interest bound. Habermas views knowledge in relation to the problems that humans face while trying to live his existence. Therefore, humans organize its experiences in terms of a *priori* interests. These knowledge-guided interests or simply called knowledge interests are classified as technical, practical and emancipatory.

Critical Research in IS

The world view of Critical research perspective focuses on the reality of interdependence of parts with the whole which necessitates that we need to study Information systems with regards to the organization, industry, and society (Orlikowski and Baroudi 1991). This philosophy attempts to critically evaluate and transform the social reality, which is in direct contrast with positivist and interpretive research philosophies. Therefore the critical researchers are not mere observers but they also try to transform the IS and organizations.

The basis of critical theory, as contended by Habermas, is to understand the activity of the subject. In *Knowledge and Human Interests*, Habermas writes that knowledge is formed in virtue of three interests: information that increases technical control; interpretations that help establish meanings; and analyses that separate consciousness from any prejudices.

Tenets

The main tenet of the critical philosophy is that social reality is historically constituted. In other words, the social actors shape the society. The critical research in IS has the following basic assumptions about the social reality (Ngwenyama 1991):

- 1) Social world including the organization and its structures are impacted by social actors.
- 2) All scientific knowledge about the social world including IS and organizations is socially constructed and thus are value laden.

- 3) Critical reasoning is used to create knowledge that will improve the organizations and individuals.
- 4) Theory and practice is combined in IS research to develop the knowledge about IS and organizations.
- 5) IS research needs to be open to critical reflection.

Knowledge Interest

The fundamental assumptions of the Critical research philosophy have been developed by Habermas (1974). In essence, Habermas believed that society would benefit more by using a critical approach to scientific knowledge claims. Knowledge interest in IS are technical, social and emancipatory (Ngwenyama 1991). Technical knowledge interest is related to human need for prediction and control of the social environment. Its focus is on building theories and methods that will benefit the organizations, IS and the users. Practical knowledge interest is related to understanding the social relations of organizations and IS. Emancipatory knowledge interest is related with getting rid of social distortions. It is concerned with improving the use of information technology (IT) in organizations by avoiding false assumptions and contradictions in use of IT.

As seen from the above classification of knowledge interest, IS has diverse interests. It is imperative that the researcher focus on both process and context from the perspective of both individual and institutional. In order to seek knowledge with critical research perspective it is necessary to adopt pluralistic inquiry methods that will help to interpret the meanings of social constructions of the inquiry.

Limitations and Criticism

Though critical theory provides an alternative perspective to study social phenomenon that will help to gain valuable insight it has also faced its share of criticisms. Karl Popper, in his article, 'Reason or revolution' (1970) wrote about critical theorists as they 'simply talk trivialities in high sounding language.' The main problem in the critical theory developed by Habermas (1974) is concerned with the relation of critique and history. Held (1980) has asked how critical theory can acknowledge its historicity and yet be critical at the same time. There seems to be a gap between theory and practice in the works of critical theorists. A difference exists in between being informed of the critical theory and actually doing critical theory research. IS researchers need to take a stance in what kind of research they want to do, whether it is critically-informed research or critical theory research. Critics say that Critical researchers are not critical enough of their own concepts. Most of the time, they delve extensively in philosophy with little time for practice.

Various authors in both the epistemology and the methodology of critical research have pointed a number of weaknesses. Although, the critical theory includes emancipation interest it lacks adequate conceptualization. Due to this inadequate conceptualization, the critical theorists emphasize on "emancipatory intent", which acknowledges that a complete emancipation is not possible. Therefore, the focus is on the

process rather than the outcome and they seek to solve social problems by promoting participation (Brooke 2002).

Publications in IS Journals

The most dominant epistemology espoused in IS research has undoubtedly been positivism. Orlikowski and Baroudi 1991 conducted a study of IS publications from 1983-1988 in four major IS research journals. Their study indicated that 96.8% of publications followed positivist approach while only 3.2 % followed interpretive approach, with no publications for critical approach.

There has been few articles in IS research that has taken critical approach of which notables are Mingers (1981), Ngwenyama (1991), Lyytinen (1992), Ngwenyama and Lee (1994), Hirschheim and Klein (1994).

An exemplar of the critical approach is the study done by Ngwenyama and Lee (1991). Authors study the communication richness in computer-mediated communication with the perspective of critical approach. They seek to develop a new theory of communication richness by offering a new definition with the help of critical perspective. The significance of this perspective is illustrated by comparing and contrasting with earlier definition of communication richness emerged from the research perspectives of positivism and interpretive. Communication richness is understood by the theory of Information richness (IRT), which states that richness of information determines whether it can change meaning over a time interval. Positivist has conceptualized the communication in IRT as a “conduit”, which acts like a medium to transport meanings from one to another. The richness of communication only depends on the choice of media in this approach. Accordingly, direct mode of communication, viz. face-to-face, is considered more superior than other modes such as email or telephone. Interpretive approach considers the ability of the sender and receiver to enact and apprehend the message. A message may have different connotations in different contexts. This fact is recognized in the interpretive approach. In essence, Positivist views communication richness as a function of channel capacity while the interpretive views it as a function of mutual understanding.

In the critical perspective, communication richness involves not only apprehending the message of the sender but also testing the validity claims of the action type of the sender. Here, action type means the type of social action taken by social actors to fulfill different goals such as: ordering to fulfill an objective (instrumental action); maintaining mutual understanding (communicative action); achieving agreement for joint action (discursive action); influencing other’s behaviors (strategic action). The definition of communication richness in critical approach goes beyond both the positivist focus on channel capacity and the interpretive focus on mutual understanding. Furthermore, the critical perspective focuses on the critique of the validity of what is being communicated and also the receiver’s emancipation of herself in case of communication noise or distortions.

Critical perspective provides us with a different view of the world than that of positivist and interpretive perspectives. As noted by Orlikowski and Baroudi (1991), critical perspective makes us aware of the interdependence of parts with the whole. Thus, while studying IS phenomenon in organizations, it should include the study of technology, organization, and the users.

Conclusion

Taking a pluralistic stance it is believed that having alternative methodologies will help to enrich the discipline since knowledge can be discovered in a certain way more fruitfully than from the other ways. Social phenomenon has intricacies and needs to be studied from different approaches. Critical approach adopts pluralistic methods of inquiry in understanding phenomenon such as participation, observation, and the analysis of contextual data. Since the technological aspect of Information Systems is rapidly changing it makes studying IS phenomenon more complex. The perspective to adopt in a research should be based on the research question, researcher's familiarity with the perspective, and the phenomenon of inquiry.

References

Contact the authors for full references

LEADERSHIP FOR THE NEXT GENERATION¹

David E. Melchar, Roger Williams University, 1 Old Ferry Rd., Bristol, RI 02809,
dmelchar@rwu.edu, 401-254-3545

Susan M. Bosco, Roger Williams University, 1 Old Ferry Rd., Bristol, RI 02809,
sbosco@rwu.edu, 401-254-3175

Collin Cantrell, Roger Williams University, 1 Old Ferry Rd., Bristol, RI 02809,
ccantrell130@hawks.rwu.edu, 401-254-3545

ABSTRACT

The technology revolution and its increasing use of electronic personal and professional media have played an instrumental role in shaping not only what interests the newest generations, but how it engages them. What leadership approach offers the best hope in the 21st Century? Human nature is not changing; it is still dependent upon self-interest and self-esteem for productive behavior in today's organizations. One leadership approach to consider is servant leadership as proposed by Robert Greenleaf [8], to see leadership as a mentoring, empowering tool, ensuring that followers have the resources and support necessary to achieve organization and personal goals.

INTRODUCTION

As four generations of workers prepare to interact in organizations for the first time, one of the important considerations is whether leadership styles that have been effective with older generations will continue to be received positively by younger workers.

Traditional leadership theories suggest that leaders have particular traits or behaviors that make them effective. Trait theory [13] states that there are six specific traits or characteristics that distinguish leaders from managers. Some of these include drive, self-confidence, and intelligence. Behavioral theories propose that leaders can be identified according to their actions. Examples of these include Lewin's [14] autocratic, democratic, and laissez-faire [2] people versus production orientation, contingency theories of leadership [7] [11] [17] and charismatic leadership [6]. Although these theories have provided explanations for leadership styles that have been successful in the past, none appears to have been able to predict which style(s) will be most effective in managing a multigenerational workforce with differing leadership expectations.

One theory that has not been adequately tested empirically is the model of servant leadership. The model of servant leadership [8] states that the leader's focus is on others rather than self and on understanding the role of the leader as servant. The servant leader takes the position of servant to his/her fellow workers, and aims to fulfill the needs of others [21]. It has been defined as serving others "by investing in their development and well being for the benefit of accomplishing tasks and goals for the common good" [15] [19]. One criticism of servant

leadership has been its lack of support from "published, well-designed, empirical research" [17, p.245] and reliance on examples that are mostly "anecdotal in nature" [17, p.245].

The literature about generational differences indicates that expectations vary among them regarding leadership characteristics. Every generation has its own expectations about the workforce, therefore, in order to lead effectively, those differences must be recognized and individuals managed accordingly [16]. Since effective leadership is essential to an organization's success, determining whether traditional models will still yield results is of the utmost importance [18]. In this paper, we will discuss leadership expectations of each generation, from the model of servant leadership, and finally the results of our study that investigated whether servant leadership is an appropriate model for all generations.

Generational Characteristics

The generations currently populating the workplace are grouped as follows: Veterans, born before 1945; Baby Boomers, born between 1945-1964; Generation X, born between 1965-1980; and Generation Y, born between 1980-2000 [4]. The percentage of workers by generation is: Veterans 10%, Baby Boomers 44%, GenXers 34%, and GenYers 12%. Leadership expectations differ among these groups. Veterans seek authority figures and are hierarchical and patriarchal [3]. Boomers are considered to be strong coaches [9]. Hastings [9] describes Gen Xers as seeking the following from their leaders: Know the reasons for decisions, encourage open debate, and engage in ongoing cooperative performance management. They also desire a mentoring relationship with a leader [5]. Gen Y workers prefer autonomy, however, they also desire close mentoring and encouragement [1] [10] [12] [16]. The two younger generations share a desire for good leadership, appreciation, feedback and the opportunity for meaningful learning [3].

Servant Leadership – A Model for All Generations

The servant leadership model can potentially change organizations and societies because it stimulates both personal and organizational metamorphoses. Empirical work by Barbuto and Wheeler [2] tested a model of servant leader characteristics wherein they developed a measure of five servant leader characteristics. These are: Altruistic calling, emotional healing, wisdom, persuasive mapping and organizational stewardship. Because its focus is on the follower rather than the leader, it should prove to be effective for all workers, regardless of generation. This leads to our main hypothesis:

H1 – There will be no generational differences in preference for the characteristics of the servant leader.

METHODOLOGY

Subjects

Subjects were recruited using a snowball sampling method to obtain a generationally representative subject pool. Of the 339 subjects, 59% were female and 41% were male. The generational breakdown was Veterans 5%, Boomers 29%, GenX 11% and GenY 55%. The actual percentage of workers in each generation is approximately 10% Veterans, 44% Baby

Boomers, 34% GenXers, and 12% GenYers, therefore, our sample had an overrepresentation of Gen Y and an underrepresentation of Gen X members. Our sample also included students, who were 77% of the GenY group, therefore, Gen Y workers accounted for only 13% of our sample.

Survey

We used a web-based survey adapted from the Barbuto and Wheeler instrument [2]. The scale contains 23 items that were measured using a 5-point Likert scale. The rater version of the survey was used and was approved by the Human Subjects Review Board of the university. There were also demographic questions included such as gender and age (to gather generational data). It was distributed online using the survey software designed by Qualtrics.

RESULTS

For the first factor of Servant Leadership, Altruistic Calling, the results of the ANOVA for generation were significant ($F=4.661, p<.01$). For Generation Y, the characteristic of Altruistic Calling, the desire to make a positive difference in the lives of followers, the mean was 2.97, significantly higher than that of any of the other generations. Those means were: Veterans 2.39, Boomers 2.38, and Gen Xers 2.56. There were no other significant differences in the means. A follow-up Tukey was conducted to determine which particular characteristics were significant. Each item in the factor was significant for different generations as seen in Table 1 below.

TABLE 1

Differences for Items within Altruistic Calling by Generation

Item Statement	F	Sig.	Generation	Mean*	Generation	Mean	P Value
This person should put my best interests ahead of his/her own	10.735	.000	Veterans	2.33	Gen Y	2.98	$p<.02$
			Boomers	2.44	Gen Y	2.98	$p<.01$
			Gen X	2.53	Gen Y	2.98	$p<.05$
This person should do everything he/she can to serve me	11.750	.000	Veterans	2.39	Gen Y	3.07	$p<.05$
			Boomers	2.41	Gen Y	3.07	$p<.01$
			Gen X	2.56	Gen Y	3.07	$p<.05$
This person should sacrifice his/her own interests to meet my needs	10.553	.000	Boomers	2.19	Gen Y	2.80	$p<.01$
			Gen X	2.42	Gen Y	2.80	$p<.10$
This person should go above and beyond the call of duty to meet my need	9.233	.000	Veterans	2.44	Gen Y	3.04	$p<.05$
			Boomers	2.49	Gen Y	3.04	$p<.01$

*Higher means indicate stronger agreement

The ANOVA for Wisdom was significant ($F=2.87, p<.05$) with means of 3.22 to 3.36. The difference occurred between the Boomers ($\bar{x}=3.22$) and the Gen Y ($\bar{x}=3.36$) respondents. The results of the follow-up Tukey (see Table 2) revealed that three of the five items were significant.

TABLE 2

Differences for Items within Wisdom by Generation

Item Statement	F	Sig.	Generation	Mean*	Generation	Mean	P Value
This person should seem alert to what's happening	2.727	.044	Gen X	3.72	Gen Y	4.10	$p<.05$
			Gen X	3.72	Boomers	4.09	$p<.10$
This person should seem in touch with what's happening	2.498	.060	Boomers	2.22	Gen Y	2.59	$p<.05$
This person should seem to know what is going to happen	3.618	.013	Boomers	3.70	Gen Y	3.97	$p<.05$

*Higher means indicate stronger agreement

For Emotional Healing, the means among generations were not significantly different. They ranged from 2.77 to 3.02. The Persuasive Mapping means were not significantly different by generation and ranged from 3.24 to 3.34. The means for Organizational Stewardship ranged from 3.79 to 3.93 with no significant differences among generations.

Therefore, Hypothesis 1 is partially supported with significant generational differences appearing for two factors - Altruistic Calling and Wisdom.

DISCUSSION AND CONCLUSIONS

The Servant Leader model has been discussed as an alternative leadership style for approximately 30 years, but has only recently been tested empirically. From the results seen here, the five items that comprise this style appear to be important to individuals, from each generation, including those who are on the brink of entering into the workplace. The characteristics that are present in a servant leader are desirable to many employees seeking a manager who is capable of being a leader as well. Perhaps it is in the aftermath of many corporate scandals of the past several years that this selfless model of leadership seems particularly appealing to so many [12]. There has been much evidence of the negative outcomes for organizations that have self-interested leadership at the helm. It seems that if a leader uses the servant leader style, he/she would be much less likely to take actions that would harm the employees and/or would harm the organization as a social entity that is part of a larger community.

Of course, as the workplace becomes more generationally diverse, it is important to consider whether the values of a servant leader are seen as equally valuable for all employees. According to our results, there are particular characteristics of servant leaders that are more salient to some generational groups than others while others were viewed similarly. Three of the factors that comprise the servant leader model were rated equally among all respondents. Emotional healing was not viewed as a particularly important factor for any of the groups. This factor had the lowest means, indicating people disagree that this is an important characteristic for a leader to possess. Perhaps the workplace is not viewed as an arena where people should expect a leader to make them feel better or be responsible for their emotional well-being. Persuasive Mapping and Organizational Stewardship were both viewed as positive characteristics for leaders to possess by all respondents. Altruistic Calling and Wisdom were viewed differently among some generational groups. The overall mean for Altruistic Calling indicated that respondents disagreed that this was a valuable leadership characteristic. Perhaps because this characteristic reflects a personal type of self-sacrifice, it is not seen by some as helpful to the organization as a whole. Wisdom was viewed positively by all generational groups overall, however, older generations rated it as less important than younger workers did.

From this study, it is apparent that most elements of servant leadership are viewed positively by respondents of all generations. With the phenomenon of four generations in the workplace just beginning, finding an effective leadership style all of them has value for organizations. It is also evident that recent corporate scandals as well as the mass downsizings that occurred 20 to 30 years ago have had different effects on each generation. There were some limitations to this study in that there was not random sampling conducted. The subjects are more homogenous than they would have been otherwise. By using a web-based instrument, the distribution to a wider pool was efficient and economical, with the resulting pool distributed adequately among the generations. However, it did necessarily limit respondents to those with access to technology. In summary, these results indicate support for the servant leadership style among all four generational groups. Considering that all employees as well as the organization benefit from effective leadership, servant leaders may find that their time is now.

REFERENCES

- [1] Barbuto, J. E., Jr., & Wheeler, D. W. (2006). Scale development and construct clarification of servant leadership. *Group & Organization Management*, 31(3), 300-326.
- [2] Blake, R.R. & Mouton, J.S. (1984). *The Managerial Grid III*. Houston, TX: Gulf Publishing.
- [3] Bland, V. (2005). Talking 'bout your generation. *The New Zealand Herald*. Retrieved October 10, 2007 from http://www.nzherald.co.nz/topic/story.cfm?c_id=229&objectid=10357066.
- [4] Burke, M.E. (2004). Generational differences survey report. *Society for Human Resource Management Research Department*. Virginia: Alexandria.

- [5] Cenek, R. (2005). From generational differences in the workplace: fact or fiction? Retrieved September 9, 2007 from http://www.energypulse.net/centers/article/article_display.cfm?a_id=1107.
- [6] Conger, J.A. & Kanungo, R.N. (2000). Charismatic leadership and follower effects. *Journal of Organizational Behavior*, 21, 747-767.
- [7] Fiedler, F.E. (1967). *A Theory of Leadership Effectiveness*. New York: McGraw-Hill.
- [8] Greenleaf, R. K. (1977). *Servant leadership: A journey into the nature of legitimate power and greatness*. New York: Paulist Press.
- [9] Hastings, R. (2007). Management skills differ between generations. Retrieved on September 6, 2007 from: http://www.shrm.org/diversity/library_published/nonIC/CMS_020430.asp.
- [10] Hastings, R.R. (2006). The Ideal Workplace for Generation Y. Retrieved September 6, 2007, from http://www.shrm.org/diversity/library_published/nonIC/CMS_019637.asp.
- [11] Hersey, P. & Blanchard, K.H. (1974). So you want to know your leadership style? *Training and Development Journal*, February, 1-15.
- [12] Kehrl, S. & Sopp, T. (2006). Managing Generation Y. *HR Magazine*, 51(5), Retrieved September 6, 2007, from <http://www.shrm.org/hrmagazine/articles/0506/0506managementtools.asp>.
- [13] Kirkpatrick, S.A. & Locke, E.A. (1991). Leadership: Do traits really matter? *Academy of Management Executive*, 5(2), 48-60.
- [14] Lewin, K. (1938). An experimental approach to the study of autocracy and democracy: A preliminary note. *Sociometry*, 1, 292-300.
- [15] Page, D. (2004). Experiential learning for servant-leadership. *Proceedings of the Servant Leadership Research Roundtable*. Retrieved May 11, 2007 from http://www.regent.edu/acad/sls/publications/journals_and_proceedings/proceedings/servant_leadership_roundtable/pdf/page-2004SL.pdf
- [16] Turetsky, D. (2006). Generations at work: New expectations & incentive requirements. *Workspan*, December. Retrieved September 13, 2007 from proquest.umi.com/helin.uri.edu.
- [17] Vroom, V.H. & Yetton, P.W. (1973). *Leadership and Decision Making*. Pittsburgh: University of Pittsburgh Press.
- [18] White, R.J. (2004). Generation integration. *Credit Union Management*, 27(10), 38-40.

¹ This project was supported by Roger Williams University through the 2008 Provost's Fund for Undergraduate Research

Routines in Emerging Organizational Structures

Arun Madapusi
University of North Texas
Department of Management, College of Business Administration
P.O. Box 305429, Denton, Texas 76203.
Phone: (940) 565 3338; E-mail: madapusi@unt.edu

Abstract

In terms of organizational structure it can be argued that practice is running ahead of theory. In trying to find ways to increase flexibility and adaptiveness and compete in a global environment, firms are restricting themselves to their core competencies and finding a variety of ways to accomplish other necessary tasks. In this research study, a routine-based approach is used to examine emerging structures as well as address how traditional organizational theory variables such as size, technology, and environment might be re-thought in terms of these emerging forms.

Keywords: Routines, Organizational Structures.

Introduction

The focus of traditional schools of thought such as neoclassical economics and resource-based views of the firm are on how to appropriate wealth once it was created [6]. Neoclassical economics suggests that firms maximize production by achieving economies of scale and use centralized structures to handle mass standardization needs [6]. The invention and diffusion initiatives that spread efficiency increases across industries are attributed to a meta-capability of coordination [2] [3]. The resource-based view suggests that firms identify and exploit unique resources and use decentralized structures to handle mass customization needs. The achievement of economies of scope is attributed to the meta-capability of delegation [2] [3]. As opposed to the above traditional views, the entrepreneurial school seeks to address the wealth creation process. The focus here is on effectively handling continuous streams of innovations through the use of inter-organizational structures such as networks and cellular forms. The knowledge sharing in these structures are attributed to the meta-capability of collaboration [2] [3].

Most studies typically focus on the impact of variables such as size, technology and environment in examining organizational issues [5] [7] [8]. There is continuing debate among researchers on the impact of these variables in the context of emerging organizational forms. Network and cellular structures are made up self-managing and autonomous teams. Size plays a significant role as smaller teams or cells are nimble and quick to capitalize on opportunities. As these cells become larger, however, a bureaucratic type structure is necessary to control and coordinate

activities, which in turn affects the operational flexibility of the cells [2] [3]. Technology is an important enabler of the network and cellular form; however, as cells become larger, firms typically use enterprise systems to garner efficiency gains, which in turn lead to a decrease in operational flexibility [2] [3]. The role of the environment is mitigated as small cells are typically fluid in form and can group and regroup to handle environmental uncertainties. As cell size increases, however, firms resort to formalized structures to coordinate activities thus surrendering flexibility for control [2] [3].

Studies indicate that a routine and predictable way of business behavior underlies much of the process changes that take place within and across firms [4] [9]. In a metaphorical sense, firms “know how to do things” and can be characterized as packages of competences. These studies suggest that the memories of individual members’ store much of the information required for the performance of organizational routines. Hence, the knowledge that a firm possesses can be reduced as well as aggregated to the knowledge of its individual members and the linking of these individual memories with shared experiences in the past [4] [9]. The above suggests that routines could provide a framework to analyze not only traditional structures but also emerging organizational forms. A routine-based approach further provides a framework to analyze the impact of variables such as size, technology, and the environment.

Researchers further indicate that often innovators tend to lose out to imitators [4] [9]. This suggests that routinizing innovations is essential for capturing first mover advantages. The “best practices” embedded in technology is captured by routines and enables firms to capture the efficiencies of the innovation. Incremental innovation is facilitated by changes to existing routines; whereas radical innovation emerges through the exploitation of anomalies in existing routines [4] [9]. The impact of cell size loses potency as across different cell sizes routines can be aggregated or de-aggregated to cater to dynamic business needs. Cells of all sizes can share routines as well as tap into and leverage the complementary routines of supply chain partners. This in turn better equips the firm to handle environmental uncertainties.

Studies indicate that a meta-capability of collaboration facilitates the process of sharing and transferring of the knowledge systems that go along with emerging structures [2] [3]. Firm competences in the form of routines underlie this meta-capability of collaboration. Increasingly organizational members are called upon to adopt more of an entrepreneurial role in organizations and the emphasis on knowledge creation feeds the process of continuous innovation. This is facilitated by frequent and repeat interactions and relationship-building within and between cells [2] [3]. The above indicates that trust plays an increasingly important role in facilitating relationship-building in emerging organizational forms [1]. This suggests that routines could underlie trust cycles in the form of tacit contracts. Thus, as new organizational forms emerge throwing up fresh challenges to organizational researchers, a routine-based framework could provide a holistic framework to tackle these issues.

References

- [1] Jap, S.D. "Perspectives on joint competitive advantages in buyer-supplier relationships." *International Journal of Research in Marketing*, 2001, 18(1/2), 19-35.
- [2] Miles, R.E., Snow, C.C., Matthews, J.A., Miles, G. *Cellular-Network Organizations* (In) W. Halal and K. Taylor (Ed.) 21st Century Economics. NY: St. Martin's Press, 1997.
- [3] Miles, G., Miles, R.E., Snow, C.C. "Wealth-Creation and Entrepreneurial Strategies: Process, Meta-Capabilities, and Barriers." *Strategic Management Society Conference*, 2000, Vancouver, British Columbia.
- [4] Nelson, R.R., Winter, S.G. 1982. *An Evolutionary Theory of Economic Change*. Cambridge, MA: Harvard University Press, 1982.
- [5] Prescott, J. E., Kohli, A. K., Venkatraman, N. "The Market-share-Profitability relationship: An empirical assessment of major assertions and contradictions." *Strategic Management Journal*, 1986, 7, 377-394.
- [6] Rumelt, R.P., Schendel, D.E., Teece, D.J. *Fundamental Issues in Strategy: A Research Agenda*. Boston: Harvard Business School Press, 1991.
- [7] Sharfman, M.P. and Dean, J.W. "Conceptualizing and Measuring the Organizational Environment: A Multidimensional Approach." *Journal of Management*, 1991, 17(4), 681-700.
- [8] Venkatraman, N., Prescott, J.E. "Environment-strategy co-alignment: An empirical test of its performance implications." *Strategic Management Journal*, 1990, 11, 1-23.
- [9] Winter, S.G. *Survival, Selection, and Inheritance in Evolutional Theories of Organization*. In J.V. Singh (ed.), *Organizational Evolution: New Directions*. Newbury Park, CA: Sage, 1990.

MODELING THE EFFECTS OF VALUE-ADDED PROCESSING AND LOGISTICS USING ARENA

Kurt A. Rosentrater, USDA, (605) 693-3241, kurt.rosentrater@ars.usda.gov
Elif Kongar, University of Bridgeport, (203) 576-4379, kongar@bridgeport.edu

ABSTRACT

Scarcity and cost of nonrenewable fossil fuels have been a growing focus of many energy-dependent nations, including the U.S., for the past couple of decades. There are mainly two ways to approach these problems: (1) reducing the energy dependency and/or (2) developing alternative means of energy production. Utilizing biofuels, which are renewable sources of energy, is a promising way to produce energy. But, as this industry grows, so too does the quantity of byproducts, which are known as DDGS. This paper discusses a model that has been developed to analyze the economic viability of DDGS processing, and includes both deterministic and probabilistic variables. Capital and operational costs for various DDGS production rates and pelleting ratios are considered in the simulation, to validate the model. The developed model is run for a total of various scenarios including different production rates and pelleting ratios. Simulation results indicate that as pelleting ratio increases, the cost to ship DDGS by rail decreases substantially. These decreases are increasingly profound as DDGS production rate increases.

Keywords: Energy, biofuels, byproducts, renewable energy, logistics, processing, simulation.

INTRODUCTION

The increasing cost of nonrenewable fossil fuels and potential decline in production in the future have been an increasing focus for many energy-dependent nations in recent years, including the U.S. There are two primary ways to approach these challenges: (1) reducing energy dependency, and/or (2) developing alternative methods of energy production.

Utilizing biofuels, which are renewable sources of energy, is one of the promising alternatives for producing energy, and can be done in a relatively efficient manner, depending on the conversion technologies that are employed. There is a variety of biomass materials that can be used to produce biofuels, viz., residue straw, corn stover, perennial grasses, legumes, and other agricultural and biological materials. Currently, the most heavily utilized material in the U.S. is corn starch, since fuel ethanol production from corn can be accomplished very efficiently and at a relatively low cost compared to other biomass sources. In fact, corn starch is currently the only biological material that can be economically converted into ethanol on an industrial scale.

The number of corn ethanol plants, and their processing capacities, has been exponentially increasing in recent years. The beginning of 2008 saw 136 manufacturing plants in the U.S. with an aggregate production capacity of 7.5 billion gal/yr (28.4 billion L/yr). In addition, 71 plants are currently under construction or expansion, and upon completion will contribute an additional 5.8 billion gal/yr (22.0

billion L/yr) [1,2]. Naturally, this growth in industry will have a direct impact on the amount of ethanol production, and thus the quantity of byproducts that are generated.

In-depth details on ethanol manufacturing, which are beyond the scope of this paper, can be found in [3-8]. Briefly, there are mainly two techniques for producing fuel ethanol using corn grain: (1) wet mill processing and (2) dry grind processing. Wet mill processing tends to be highly capital intensive, whereas dry grind processing has significantly lower capital and operational requirements. Therefore, dry grind processing has rapidly gained prevalence in the industry.

The dry grind production process (Figure 1a) consists of several key steps, including grinding, cooking, liquefying, saccharifying, fermenting, and distilling. Typically, ethanol is the primary end product of this process along with two byproducts: residual nonfermentable corn kernel components and carbon dioxide.

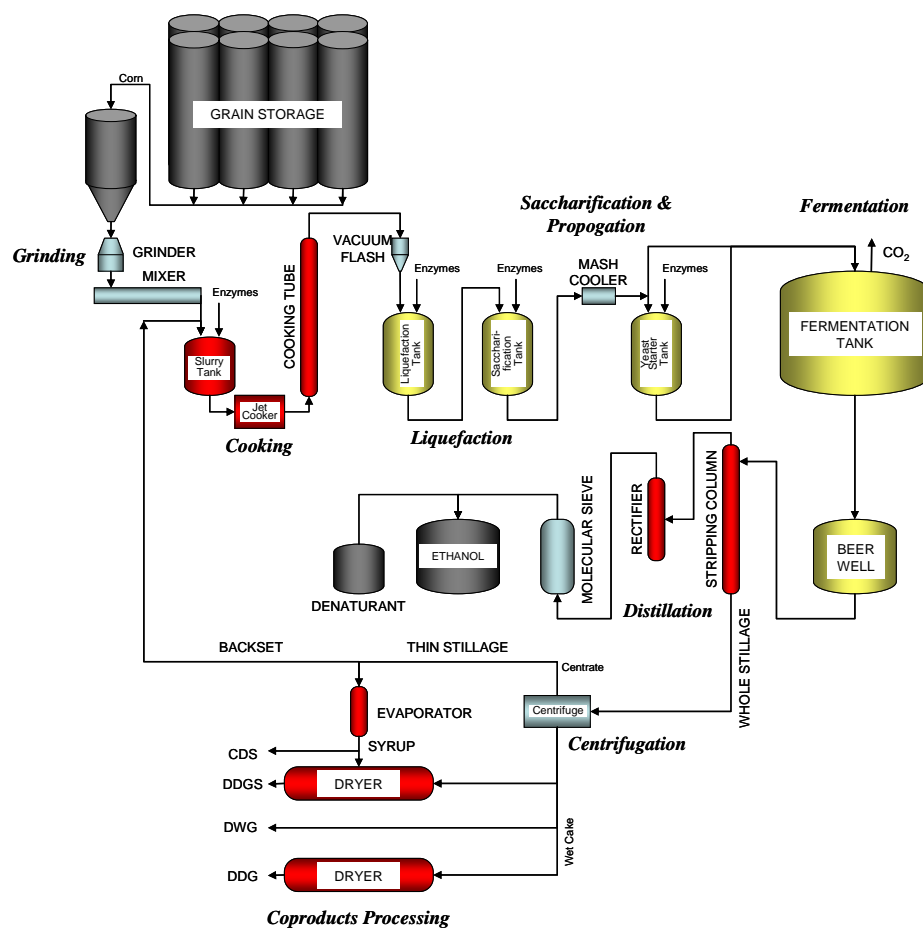


Figure 1a. Process Flowchart for Dry Grind Fuel Ethanol Production.

The nonfermentable kernel components (i.e., corn protein, fiber, and oil) are usually further processed (Figure 1b) and then marketed in the form of distillers dried grains with solubles (DDGS), and to a lesser degree in the form of distillers dried grains (DDG), which do not contain added solubles, distillers

wet grains (DWG), and condensed distillers solubles (CDS). Hereafter “distillers grains” will be used in a generic sense to refer to all of these byproduct materials.

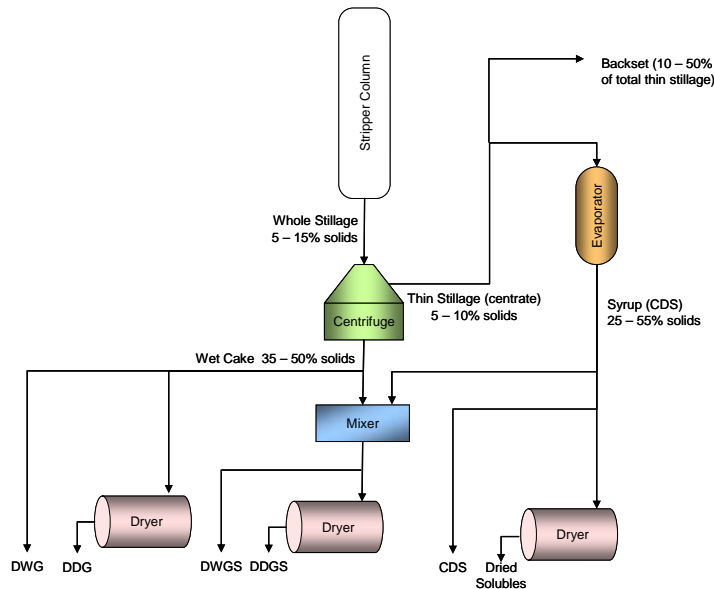


Figure 1b. Process Flowchart for Ethanol Byproduct Processing.

Residue streams are separated from the ethanol during distillation. They are often dried to approximately 10% moisture content, to ensure a long shelf life and facilitate transportation, and are then sold as distillers grains (most often DDGS or DDG) to local livestock producers or shipped via rail to livestock feed markets (Figure 2) throughout North America, and they are increasingly being exported to overseas destinations as well.



Figure 2. Simplified Flowchart of Current DDGS Processing Steps that Could be Augmented by Value-Added Processing.

The sale of distillers grains contributes substantially to the economic viability of ethanol manufacturing (generally between 10 and 40% of an ethanol plant’s entire revenue stream, depending upon DDGS sales price and other market conditions), and is thus a vital component to each plant’s operations. Because of the dynamics of the free market economy under which this industry operates, the quantity of processing residues that will be produced as this industry continues to grow (and the ability to utilize them) will significantly impact the future of the industry.

A persistent barrier to effective utilization of distillers grains is product storability and flowability. DDGS is a granular, powder-like material that consists of a range of particle sizes and shapes. DDGS is typically shipped in trains and trucks throughout the U.S. to be used for animal feed. However, DDGS is often difficult to unload once the vessel reaches destination, because the particles lock together and have flowability problems. This necessitates manual unloading processes, which creates a substantial financial burden for the ethanol manufacturer. This flowability problem is resulting in serious economic implications for ethanol plants (e.g., sledgehammers, shovels, and pick axes must be used to get the DDGS to discharge). These economic losses include rail car repairs (due to damage), and labor expenses.

Another issue with the transportation of DDGS is the weight vs. volume of each car. Railcars containing DDGS are filled to volumetric capacity for shipping, but are often not at maximum allowable weight, due to the low bulk density of the granular material itself (approximately 30 lb/ft³ [480 kg/m³]), thus causing additional potential economic loss to the ethanol manufacturer.

Thus there exists a serious need to modify the DDGS that is currently produced in the industry (and inherently the processes which are used to produce the DDGS), or to alter the storage and transportation systems for these products (which could be a very expensive proposition).

Pelleting is a manufacturing process that is commonly used to densify granular materials. This process is both feasible and appropriate for DDGS, and can be accomplished using conventional feed milling equipment [9]. In fact, pelleting of DDGS (Figure 3) would simultaneously overcome the flowability problem, as well as allow shipping vessels to be filled to both weight and volumetric capacity. Ultimately, however, the adoption of this technology is contingent upon price ramifications to the ethanol plant, which has not yet been thoroughly explored.

Cost-effectiveness of pelleting DDGS versus current industrial practice, which is to ship DDGS as-is (i.e., in granular bulk form) has to be investigated. Rosentrater (2007) [9] has proven the feasibility of this type of processing, and has shown that DDGS can be pelleted.

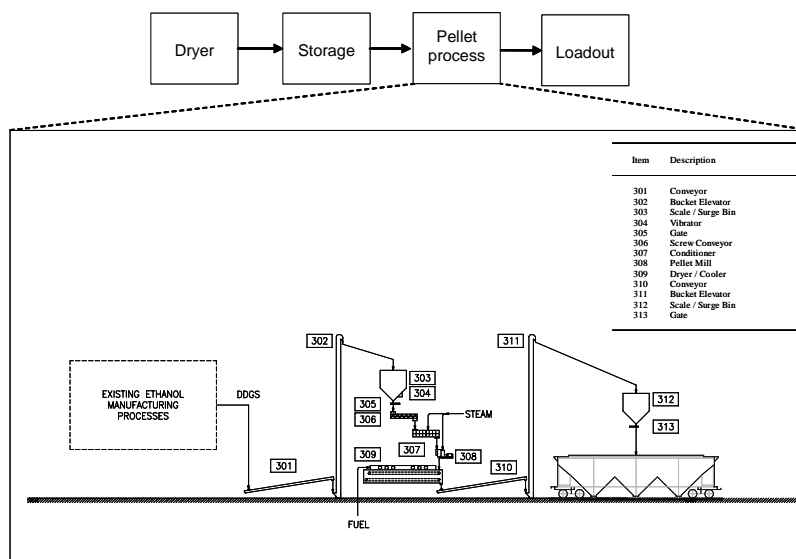


Figure 3. Proposed Flowchart for DDGS Processing Augmented by Pelleting.

The question that remains, however, regards the cost of this additional processing step. In order to answer this question, a computer model was developed and implemented to examine the economic implications of pelleting DDGS for varying production and pelleting rates with the help of simulation models. The computer models were developed using Arena v. 10.00.00 – CPR 7 (Rockwell Software, Inc.).

2. DETERMINISTIC AND STOCHASTIC SIMULATION MODELS FOR DDGS PELLETING

In order to investigate the economic implications of pelleting DDGS, two separate computer models were developed. The first simulation model employed only deterministic variables, whereas the second model embodied stochastic variables to take possible dynamic variations into account. The deterministic model discussed in this paper was previously described by Rosentrater and Kongar [14]. The following discussion briefly explains the algorithms that were used for both the deterministic and stochastic simulation models.

2.1. Algorithm Development

Both dependent and independent variables were embedded in the simulation models as follows. These were based on data available for a range of current ethanol processing facilities.

i. Data for daily DDGS generation, g , (tons/day) are given as in (1):

$$g = 0, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000. \quad (1)$$

Data for the percentage of DDGS pelleted (%) are given as in (2):

$$p = 0, 25, 50, 75, 100. \quad (2)$$

ii. Bulk density of p percent pelleted DDGS (tons/ft³), d_p , can be defined as in (3.0) for the deterministic model:

$$\begin{aligned} d_0 &= 30 \text{ lb/ft}^3 = 0.0134 \text{ tons/ft}^3 \\ d_{25} &= 32.5 \text{ lb/ft}^3 = 0.0145 \text{ tons/ft}^3 \\ d_{50} &= 35 \text{ lb/ft}^3 = 0.0156 \text{ tons/ft}^3 \\ d_{75} &= 37.5 \text{ lb/ft}^3 = 0.0167 \text{ tons/ft}^3 \\ d_{100} &= 40 \text{ lb/ft}^3 = 0.0179 \text{ tons/ft}^3. \end{aligned} \quad (3.1)$$

As for the stochastic model, it is estimated that the bulk density will vary by +/- 5% of above stated values (each level of pelleting has its own d_p value), which will be distributed normally. Hence,

$$\begin{aligned} d_o &= |N(30,0.01)| \text{ lb/ft}^3 \\ d_{25} &= |N(32.5,0.01)| \text{ lb/ft}^3 \\ d_{50} &= |N(35,0.01)| \text{ lb/ft}^3 \\ d_{75} &= |N(37.5,0.01)| \text{ lb/ft}^3 \\ d_{100} &= |N(40,0.01)| \text{ lb/ft}^3, \end{aligned} \quad (3.2)$$

Where, $|\mu, \delta|$ is the absolute value of the normally distributed variable with parameters μ and δ .

iii. Pelleted quantity, pq , (ton/day) can be calculated as in (4):

$$pq = g^*p. \quad (4)$$

The total operating cost of pellet processing can be given as in Table 1 for the deterministic model.

Table 1. Total Operating Cost of Pelleting Process for the Deterministic Simulation Model

p (tons)	Co_p (\$/ton)	p (tons)	Co_p (\$/ton)
100	1.4700	600	1.1025
200	1.3965	700	1.0290
300	1.3230	800	0.9555
400	1.2495	900	0.8820
500	1.1760	1000	0.8085

As for the stochastic model, the values for Co_p are estimated to vary by +/- 5% of the values stated in Table 1 (i.e., each level of pelleting has its own Co_p value, as shown in Table 2), and will be distributed normally.

Table 2. Total Operating Cost of Pelleting Process for the Stochastic Simulation Model

p (tons)	Co_p (\$/ton)	p (tons)	Co_p (\$/ton)
100	N(1.4700,0.05)	600	N(1.1025,0.05)
200	N(1.3965,0.05)	700	N(1.0290,0.05)
300	N(1.3230,0.05)	800	N(0.9555,0.05)
400	N(1.2495,0.05)	900	N(0.8820,0.05)
500	N(1.1760,0.05)	1000	N(0.8085,0.05)

iv. Transportation costs

Transportation variables, along with their corresponding descriptions and values, are provided in Table 3 for the deterministic model.

Table 3. Shipping Cost Data for the Deterministic Simulation Model

<i>Variable</i>	<i>Description</i>	<i>Value</i>
x	DDGS shipping cost per rail car	\$4,800/car
y	Maximum volumetric capacity per rail car	6,350 ft ³ /car
wt	Maximum theoretical weight per rail car	109 ton/car
s	DDGS sales price	\$120/ton
t	Train capacity	10,447.5 tons
	Maximum cars/unit train	96

As for the stochastic model, it is estimated that the DDGS shipping cost per rail car (x) will follow a triangular distribution with a minimum value of 4,000 \$/car, maximum value of 5,000, and a most likely value of 6,000. In addition, DDGS sales price is also assumed to follow a triangular distribution with a minimum value of 61.6 \$/ton, maximum value of 131, and a most likely value of 81 (Table 4). These parameters were obtained from the historical DDGS sales price data depicted in Figure 4.

Table 4. Shipping Cost Data for the Stochastic Simulation Model

<i>Variable</i>	<i>Description</i>	<i>Value</i>
x	DDGS shipping cost per rail car	TRI(4000;5000;6000) \$/car
y	Maximum volumetric capacity per rail car	6,350 ft ³ /car
wt	Maximum theoretical weight per rail car	109 ton/car
s	DDGS sales price	\$120/ton
t	Train capacity	10,447.5 tons
	Maximum cars/unit train	96

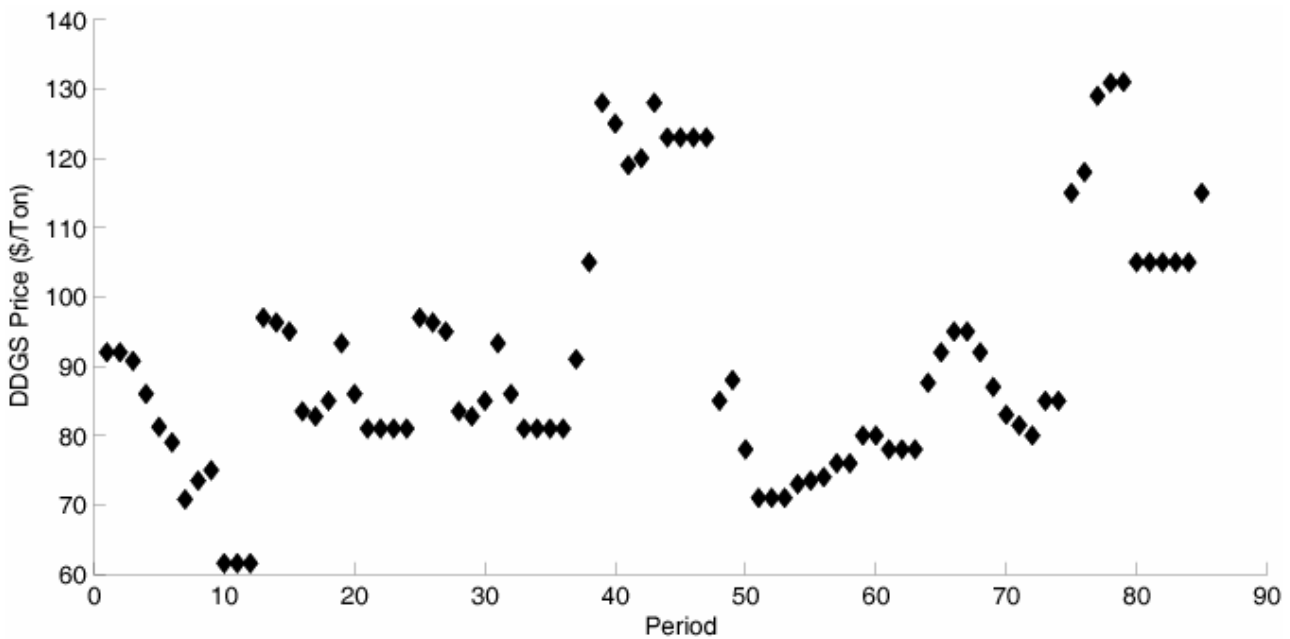


Figure 4. Monthly Data for the DDGS Price Between October 2000 and October 2007 [15].

Depending on the data provided in Table 3 and Table 4, the actual car weight when filled according to *per percent* pelleted DDGS in (ton/car) can be expressed mathematically as in (5).

$$wa_p = y * d_p \tag{5}$$

Theoretical shipping cost of *per percent* pelleted DDGS in (\$/ton) can be given as:

$$Ct_p = x / wt, \tag{6}$$

where, actual shipping cost of *per percent* pelleted DDGS in (\$/ton) is:

$$Ca_p = x/wa_p. \quad (7)$$

Hence, slack (unused) capacity in (\$/ton) can be calculated as:

$$w_p = Ca_p - Ct_p \quad (8)$$

Depending on Eq. (8) total slack cost per car can be calculated as in (9).

$$SC_{car} = s(wt - wa_p) + w_p(wt - wa_p) \quad (9)$$

Where w_p is the shipping expense for unused space in the rail car. Pelleting cost per car (P_{car}) and net cost per car (N_{car}) be calculated as in (10) and (11):

$$P_{car} = (Co_p - wa_p \cdot p/100), \quad (10)$$

$$N_{car} = SC_{car} + P_{car} . \quad (11)$$

Theoretical number of cars can be given as:

$$rt = \lceil t/wt \rceil, \quad (12)$$

where function $\lceil \cdot \rceil$ returns the nearest larger integer. Actual number of cars can be expressed mathematically as:

$$ra = \lceil t/wa_p \rceil. \quad (13)$$

Theoretical leasing fee (lt) and actual leasing fee (la) can be given as in (14) and (15):

$$lt = rt * x, \quad (14)$$

$$la = ra * x. \quad (15)$$

Therefore, the rail car leasing slack value in (\$/train) is:

$$ls = la - lt \quad (16)$$

Depending on above equations, the total pelleting cost (P) in (\$/train) and the net cost impact (N) in (\$/train) can be given as below:

$$P = ra * P_{car}, \quad (17)$$

$$N = P + ls. \quad (18)$$

The number of trains required per month can then be calculated as:

$$nt = 1(t/g) * 30, \quad (19)$$

whereas, the net cost impact (NCI) to the ethanol plant can be expressed mathematically in (\$/month) as in (19):

$$NCI = N*nt. \quad (20)$$

RESULTS AND DISCUSSIONS

Results for the net cost per car to ship DDGS (Figure 5.1.a) indicate that, compared to the baseline case (i.e., shipping DDGS as-is, with no pelleting), all scenarios that involve pelleting have reduced shipping costs. And, the least shipping costs are produced when 100% of the DDGS is pelleted. As DDGS generation rate increases, a slight decrease in shipping cost is observed (due to economies of scale). On the other hand, when the fraction pelleted increases, large decreases in shipping costs are observed (due to the weight versus capacity for each car). For example, at a DDGS generation rate of 100 tons/day, 0% pelleted DDGS had a shipping cost 19.8 times greater than 100% pelleted DDGS. At 1,000 tons/day, the 0% pelleted had a cost 36.0 times greater than 100% pelleted.

Results for the net cost per train to ship DDGS (Figure 5.2.a) once again indicate that, compared to the baseline case (i.e., no pelleting), all scenarios that involve pelleting have reduced shipping costs. The least shipping costs are produced when 100% of the DDGS is pelleted. As DDGS generation rate increases, a slight decrease in shipping cost is observed (due to economies of scale). On the other hand, when the fraction pelleted increases, large decreases in shipping costs are observed. For example, at a DDGS generation rate of 100 tons/day, 0% pelleted DDGS had a shipping cost 8.4 times greater than 100% pelleted DDGS. At 1,000 tons/day, though, the 0% pelleted had a cost 15.3 times greater than 100% pelleted.

As expected due to the previous results [14], the net cost impact per month to ship DDGS (Figure 5.3.a) indicates that, compared to the baseline case (i.e., no pelleting), all scenarios that involve pelleting have reduced shipping costs. And, the least shipping costs are produced when 100% of the DDGS is pelleted. As both DDGS generation rate and the fraction pelleted increase, large increases in monthly cost impact are produced (because of overall decreases in shipping cost as the fraction pelleted increases, which is due to both economies of scale as well as weight versus capacity for each car). For example, at a DDGS generation rate of 100 tons/day, 0% pelleted DDGS had a shipping cost 8.4 times greater than 100% pelleted DDGS. At 1,000 tons/day, the 0% pelleted had a cost 15.3 times greater than 100% pelleted.

As for the stochastic results, the baseline case always seems to be dominated by pelleting options. In addition, the introduction of stochastic variables seem to decrease the net cost per car to ship (4.2.a). However, for both net cost per train to ship and the net cost impact per month to ship DDGS stochastic variables seem to significantly increase the overall cost values compared to the deterministic results. These results are probably more reflective of real life, as they more realistically depict the dynamic market place in which the fuel ethanol industry operates. The next step for this line of research should include model verification and validation.

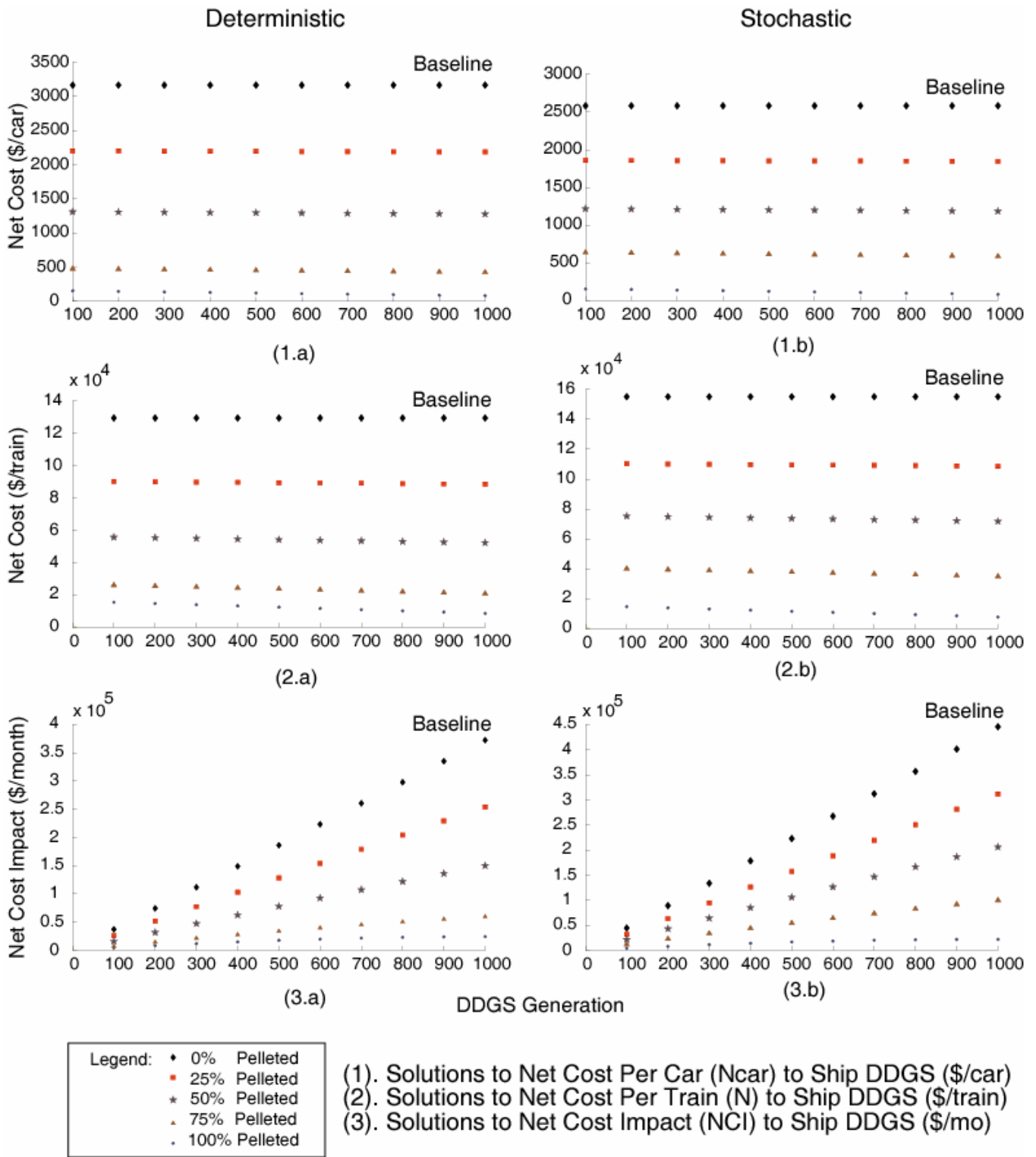


Figure 5. Deterministic vs. Stochastic Simulation Results.

CONCLUSIONS

In this paper both deterministic and stochastic simulation models are presented to investigate the economics of a proposed pelleting process for DDGS. The focus of this work was to compare two methods of utilizing ethanol byproducts: (1) value-added processing of DDGS by pelleting, and (2) current industrial practice, which is to ship DDGS as-is (i.e., in granular bulk form). The simulation results indicate that pelleting is indeed cost-effective for these manufacturing residues. In terms of cost/car and cost/train, pelleting has been shown to reduce costs for shipping DDGS for all production rates studied – the greater the fraction pelleted, the greater the cost savings. In terms of cost/month to an ethanol plant, it appears that both production rate and fraction pelleted have considerable effects, all of which point toward a greater cost savings when the DDGS is completely pelleted prior to shipping.

REFERENCES

- [1] BBI, 2008. U.S. Production Capacity. Existing Plants. BBI International. Available online: <http://www.ethanolproducer.com/plant-list.jsp>. [Accessed 08 January 2008].
- [2] RFA, 2008. Ethanol industry outlook 2007. Washington, D.C.: Renewable Fuels Association. Available online: <http://www.ethanolrfa.org/industry/locations/>. [Accessed 08 January 2008].
- [3] Bothast, R.J. and Schlicher, M.A., 2005. Biotechnological processes for conversion of corn into ethanol. *Applied Microbiology and Biotechnology* 67: 19-25.
- [4] Dien, B.S., Bothast, R.J., Nichols, N.N., Cotta M.A., 2003. The U.S. corn ethanol industry: an overview of current technology and future prospects, in: *The Third International Starch Technology Conference – Coproducts Program Proceedings*, eds. M. Tumbleson, V. Singh, and K. Rausch, 2-4 June, 2003, University of Illinois, pp. 10-21.
- [5] Jaques, K.A., Lyons, T.P., Kelsall D.R., 2003. *The Alcohol Textbook*. Nottingham University Press: Nottingham, UK.
- [6] Maisch, W.F., 2003. Fermentation processes and products, in: *Corn Chemistry and Technology*, 2nd ed. White, P.J. and Johnson, L.A. eds. American Association of Cereal Chemists: St Paul, MN.
- [7] Tibelius, C., 1996. Coproducts and Near Coproducts of Fuel Ethanol Fermentation from Grain. Agriculture and Agri-Food Canada – Canadian Green Plan Ethanol Program: Starchy Waste Streams Evaluation Project. Available online: http://res2.agr.ca/publications/cfar/index_e.htm. [Accessed 07 January 2007].
- [8] Weigel, J.C., Loy, D., and Kilmer, L., 2005. Feed Co-Products of the Dry Corn Milling Process. Iowa State University and Iowa Corn Promotion Board. Available online: www.iowacorn.org/ethanol/ethanol_17.html. [Accessed 07 January 2007].
- [9] Rausch, K.D., Belyea, R.L., 2006. The future of coproducts from corn processing. *Applied Biochemistry and Biotechnology* 128, 47-86.
- [10] Rosentrater, K.A., Giglio, M., 2005. What are the challenges and opportunities for utilizing distillers grains? *Distillers Grains Quarterly* 1(1), 15-17.
- [11] Rosentrater, K.A., 2006. Expanding the role of systems modeling: considering byproduct generation from biofuel production. *Ecology and Society* 11(1), 1-12.

- [12] UMN, 2008. The Value and Use of Distillers Dried Grains with Solubles (DDGS) in Livestock and Poultry Feeds. University of Minnesota. Available online: <http://www.ddgs.umn.edu/>. [Accessed 08 January 2008].
- [13] Rosentrater, K. A., 2007. Can you really pellet distillers grains? *Distillers Grains Quarterly* 3(3): 16-21.
- [14] Rosentrater, K., Kongar, E., 2007. "Biofuel Coproducts: Modeling the Effects of Value-Added Processing and Logistics", 5th International Logistics and Supply Chain Congress 2007, November 7-9, Istanbul, Turkey.
- [15] USDA ERS, 2008. Feed Grains Database: Yearbook Tables. United States Department of Agriculture, Economic Research Service. Available online: www.ers.usda.gov/data/feedgrains/. [Accessed 07 January 2008].

AN ESTIMATION AND ASSIGNMENT OF CO2 EMISSIONS IN LOGISTICS

Tadayuki Masui, Masayuki Goto, Tetsuo Yamada*, Tomokazu Yoshifuji
Faculty of Environmental and Information Studies, Musashi Institute of Technology
Yokohama, Kanagawa, Japan, tyamada@yc.musashi-tech.ac.jp, (+)81-45-910-2916
(*Corresponding author)

ABSTRACT

To prevent global warming, CO2 emissions should be reduced throughout supply chain. Logistics among manufacturers, distributors and consumers for products still have high environmental loads for CO2 emissions. However, issues still remain regarding how to grasp and estimate CO2 emissions for each product and how to determine an assignment to owners of products. This study focuses on a logistics process for fresh vegetables and frozen foods in a Japanese CO-OP case, and presents an estimation and assignment method of CO2 emissions. First, logistics processes in the CO-OP case are surveyed, and each workflow such as unloading, storage and loading is defined among manufacturers, distributors and consumers. Next, the estimation method of CO2 volumes is presented for each process. Finally, the survey results show the differences of the grasp and estimation method between the two products and compares CO2 volumes between them numerically.

Keywords: Environmental Logistics, Logistics Process, Logistics Workflow, Supply Chain

INTRODUCTION

Now CO2 emissions must be reduced throughout the whole product lifecycle as soon as possible. Even in Japan, logistics among manufacturers, distributors and consumers for products still have high environmental loads for CO2 emissions [2]. According to a questionnaire survey in a Japanese CO-OP (Co-operative Union) case [5], most consumers have high environmental consciousness, and the reduction of CO2 emissions strongly depends on whether industries can propose a better business model and framework with environmental logistics in practice or not. However, issues still remain regarding how to grasp and estimate CO2 emissions for each product and how to determine an assignment to owners of products for CO2 volumes of each product [1][2].

This study focuses on a logistics process for frozen foods and fresh vegetables in a Japanese CO-OP case, and presents an estimation and assignment method of CO2 emissions. First, logistics processes in the CO-OP case are surveyed based on [4] [5], and each workflow such as unloading, inspection, storage and loading is defined among manufacturers, distributors and consumers. Next, the estimation method of CO2 volumes is presented for each process. Finally, the survey results show the differences of the grasp and estimation method between the fresh vegetables and frozen foods, and the CO2 volumes between them are numerically compared.

SURVEY METHOD

Table 1 shows a survey overview for logistics processes in a Japanese CO-OP case. Japanese spinach as a fresh vegetable and Japanese chicken thighs as a frozen food are surveyed and compared. Since the survey is different between the two products as shown in Table 1, the workflows in both the logistics

processes are defined and compared, and CO₂ volumes per product are finally estimated. The logistics process consists of Procurement, Main and Home Delivery Logistics, and the procurement has two functions (Farmers / Manufacturers and Collection Terminal / Warehouse).

Table 1 A Survey Overview for Logistics Processes in a Japanese CO-OP case

Products	Type	Fresh Vegetables	Frozen Foods
	Name	Komatsuna (Japanese Spinach)	Japanese Chicken Thighs
	Unit Weight	200 [g]	415 [g]
	Unit Cubic Volume	-	1.33 [l]
Period	Survey	Aug 2005 to Jan 2006	Dec 2006 to Feb 2007
		6 Months	3 Months
	Used Data	Apr 2005 to June 2005	Oct 2006
Logistics Process		3 Months	1 Week
	Procurement	Farmers	(Manufacturers)
		Collection Terminal	Warehouse
	Main	Setting Center	Setting Center
	Home Delivery	2 Deliverly Centers	(1 Deliverly Center)

Note: () indicates no survey was conducted in this study.

DEFINITION OF WORKFLOW IN A LOGISTICS PROCESS

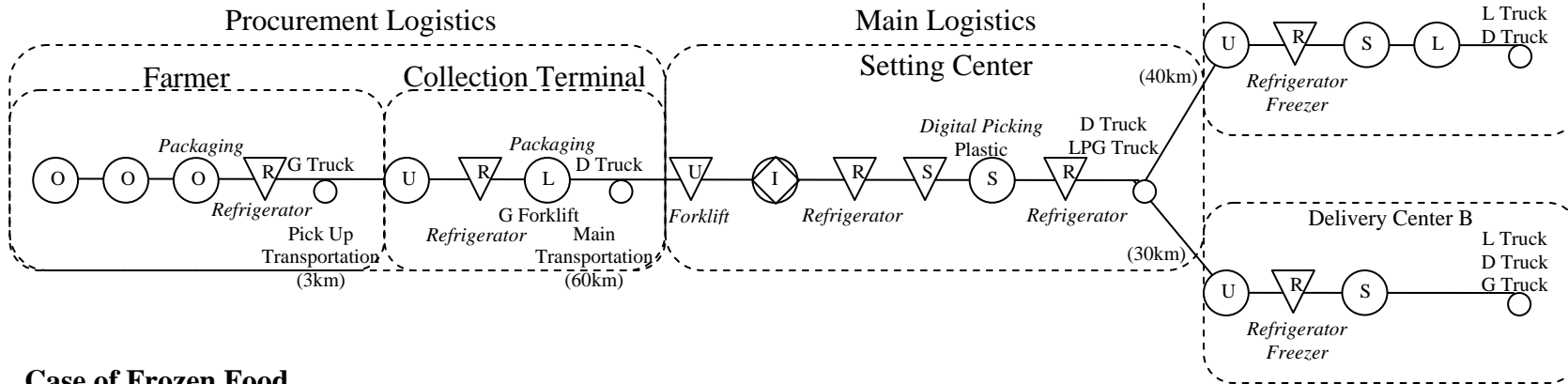
The workflows are defined and compared in each logistics process. Figure 1 shows the defined workflow for logistics processes of fresh vegetables and frozen foods in the case of a Japanese CO-OP. In view of types of workflows, both products have basically similar workflows and sequences, such as unloading, storage, loading and transportation. For storage, freezers with higher electric consumptions are used for the frozen foods instead of refrigerators which are used for the fresh vegetables.

In the procurement logistics process, products are first transported by diesel or gasoline truck as a pick up transportation from the farmers to the collection terminal in the case of fresh vegetables or from the manufactures to the warehouse in the case of frozen food. At the collection terminal / warehouse, the products are unloaded by electric-powered forklifts or manual handlings, the fresh vegetables are kept in refrigerators and the frozen food in freezers. After that, the products are shipped by diesel truck to the next setting centers as a main transportation. In the main logistics process, the inspection and setting / sorting are carried out as well as unloading, storage and loading. In the home delivery logistics process, only the case of the fresh vegetables is surveyed. The vegetables are delivered to consumers by gasoline, diesel and LPG trucks.

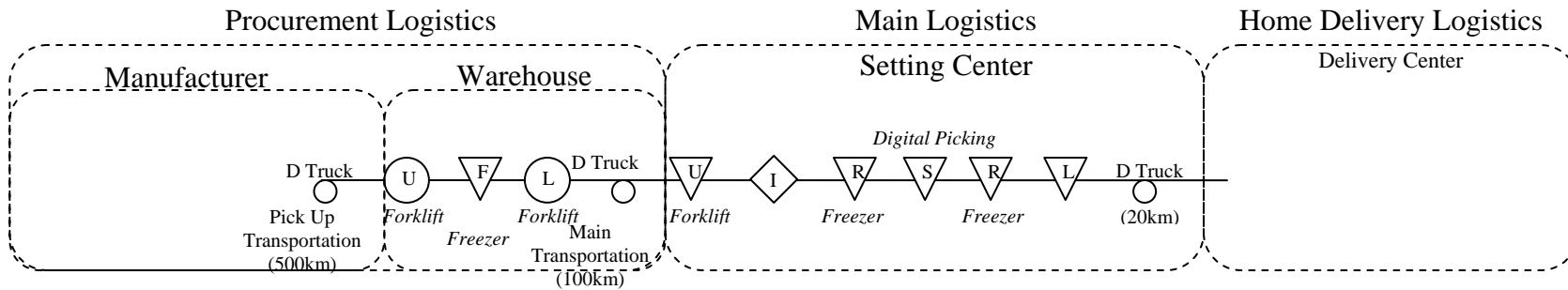
AN ESTIMATION AND ASSIGNTMENT MOTHOD OF CO₂ VOLUMES

Based on the defined workflows (Figure 1), an estimation and assignment method of CO₂ volumes are proposed for each work process as shown in Tables 2 and 3. The total estimated CO₂ volumes during the surveyed period are basically calculated by multiplying the energy / material consumptions by the CO₂ emission coefficient. However, it is difficult for real cases to obtain all actual data of the consumption volumes. The assigned CO₂ volumes per product are obtained by dividing the total estimated CO₂ volumes by the appropriate number of products in each piece of equipment / material. Ideally, the appropriate number of products should be calculated by unit cubic capacity for a refrigerator and freezer, unit weight for a forklift and truck.

Case of Fresh Vegetable
(Japanese Spinach)



Case of Frozen Food
(Fried Chicken)



- ▽ U : Unloading
- ▽ R, ▽ F : Storage by Refrigerator or Freezer
- L : Loading
- : Transportation
- ◇ I, ◇ I : Inspection
- ▽ S, ○ S : Setting / Sorting
- O : Others
- G Truck : Gasoline
- D Truck : Diesel
- LPG Truck : LPG Gas

Note: *Italic* means electric power.

Figure 1 Defined Workflows in Each Logistics Process: a Japanese CO-OP Case

Table 2 The estimation and assignment method for each work process for the Collection Terminal / Warehouse in Procurement Logistic

Types of energy / materials	Workflow process	Equipment / materials	Grapsed total energy / material consumption		Divided by a number for assignment	
			Fresh vegetables	Frozen foods	Fresh vegetables	Frozen foods
Gasoline / light fuel	Pick up transportation	Gasoline / diesel Truck	estimated travel distance × number of trips × fuel consumption rate [l/km]	travel distance per day × number of business days × fuel consumption rate [l/km]	number of products shipped [units]	truckload quantity / unit weight for chicken
Gasoline	Unloading/ loading	Forklift	actual replenishment quantity × fuel consumption rate [l/km]	-	total / unit weight for spinach	-
Electric power	Storage	Refrigerator / freezer	estimated by power rating and busy period	estimated by power rating , busy period and power coefficient	container capacity for all products / container capacity for spinach	(each cubic capacity × each average inventory day × number of products for all products) / unit cubic capacity × average inventory day for chicken
Light fuel	Main transportation	Diesel truck	estimated travel distance [km] × fuel consumption rate [l/km]	travel distance per day × number of business days × fuel consumption rate [l/km]	total / unit weight for spinach	truckload quantity / unit weight for chicken

Table 3 The estimation and assignment method for each work process in Main Logistics

Types of energy / materials	Workflow process	Equipment / materials	Grapsed total energy / material consumption		Divided by a number for assignment	
			Fresh vegetables	Frozen foods	Fresh vegetables	Frozen foods
Electric power	Unloading/ loading	Forklift	estimated by power rating, busy period and number of equipments	estimated by power rating and busy period	products processed [units]	total / unit weight for chicken
	Storage	Refrigerator / freezer	actual volume of electric power consumption	estimated by power rating , busy period and power coefficient	number of products processed [units]	(each cubic capacity × number of products for all products) / unit cubic capacity
	Setting / sorting	Digital picking	actual volume of electric power consumption	estimated by power rating , busy period and power coefficient	number of products shipped [units]	weight for number of products shipped
Light fuel / LPG	Transportation	Diesel / LPG truck	actual volume of electric power consumption	-	number of products processed [units]	-
			trial travel distance [km] × fuel consumption rate [l/km] × number of outward trips × (1+2/3)	travel distance per day × number of business days × fuel consumption rate [l/km]	the CO2 volumes [kg-CO2] / number of products processed [units]	truckload quantity / unit weight for chicken

In the case of the frozen foods, these calculations are applied as much as possible. Similar to the calculation of the estimated CO2 volumes, it is difficult for real cases to obtain all actual data of the consumptions. In such a case, the estimated consumptions are calculated as the number of products processed or shipped.

NUMERICAL CONSIDERATIONS FOR ESTIMATED AND ASSIGNED CO2 EMISSTION

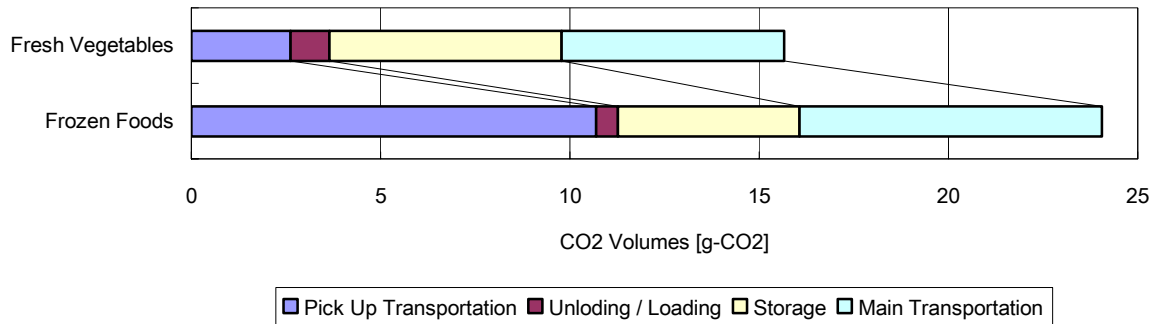


Figure 2 Estimated and Assigned CO2 Volumes per Product in Procurement Logistics Process

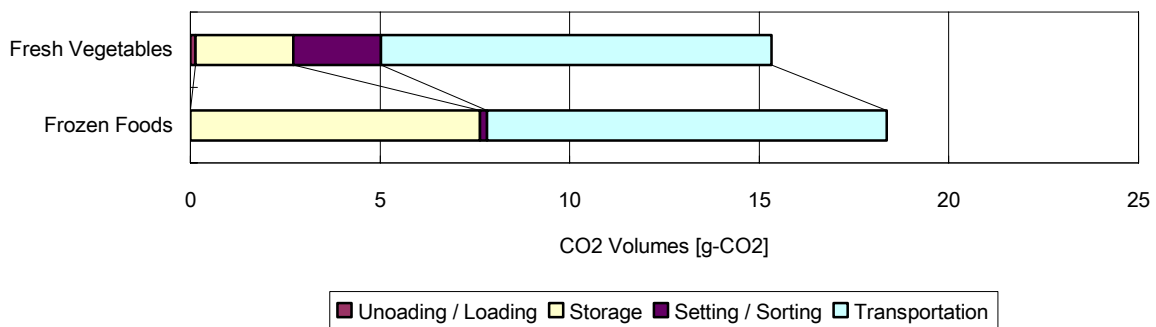


Figure 3 Estimated and Assigned CO2 Volumes per Product in Main Logistics Process

Figure 2 shows the estimated and assigned CO2 volumes per product in the procurement logistics process. For each product, the pick up and main transportation processes have a higher CO2 volume than any of the other processes. In particular, the volume of the pick up transportation process for the frozen food is highest throughout the procurement and main logistics processes, and a distance between the manufacturer and warehouse is the longest in all transportation processes surveyed. Therefore, the improvement of the transportation process should be most effective in procurement logistics.

Figure 3 shows the CO2 volumes per product in the main logistics process. Similar to the procurement logistics process, the transportation process has higher CO2 volumes than any other process for each product. For storage, the CO2 volume of the frozen food is more than three times that of fresh vegetables. This is the case because the quality and freshness of the frozen food seldom decreases due to the number of inventory days. Some of the frozen food items remain in the freezer for more than 6 months. Therefore, the inventory control with the FIFO rules for shipment and the demand-to-supply management should be strictly operated to reduce the CO2 volumes. For setting and sorting, the CO2 volume of the fresh vegetables is more than twelve times that of frozen foods. This is because each spinach is individually packaged by a plastic, and the packaging machine consumes the amount of

electric power. It might be acceptable for most environmentally-conscious consumers even if the packaging process is simplified.

Table 4 shows the total CO2 volume per product thorough the whole logistics process surveyed. In view of the sub total volume, there is a small difference in the main logistics, however, the total CO2 volume of fresh vegetables is lower than that of frozen foods by 37% in the whole logistics processes. The difference for pick up transportation in the procurement logistics directly impacts on the total volume.

Table 4 Total CO2 Volume per Product in the Whole Logistics Processes Surveyed

Types of products	Logistics process	Unloading / Loading	Storage	Setting / Sorting	Pick up transportation	(Main) Transportaion	Sub total CO2 volume per prudoct [g-CO2]	Total CO2 volume per product [g-CO2]
Fresh Vegetables	Procurement	1.03	6.13	-	2.62	5.88	15.66	30.98
	Main	0.13	2.59	2.31	-	10.30	15.32	
Frozen Foods	Procurement	0.58	4.80	-	10.69	7.99	24.06	42.42
	Main	N/A	7.63	0.19	-	10.54	18.36	

SUMMARY AND FUTURE STUDIES

This study focuses on a logistics process for frozen foods and fresh vegetables in a Japanese CO-OP case, and presents an estimation method and assignment of CO2 emissions. By the proposed method, CO2 volume per product throughout the logistics processes is grasped and numerically shown for the two products. According to the results, the total CO2 volume of fresh vegetables is lower than that of frozen foods by 37% because of a long distance in pick up transportation for frozen foods, and it is also shown that a simple packaging for fresh vegetables and the inventory control in freezers for frozen foods might contribute to reduce the total CO2 volumes. Further studies should develop the estimated and assigned method for the other types of products, propose a way to grasp data accurately for energy consumption, improve the work policies for lower environmental loads, etc.

ACKNOWLEDGMENTS

The CO-OP are gratefully appreciated for cooperating in this survey. This research was partially supported by the Japan Society for the Promotion of Science (JSPS), a Grant-in-Aid for Scientific Research (B), 19310100 in 2007.

REFERENCES

- [1] Environmental Logistics Survey Committee. *Report of Environmental Logistics Survey*. Japan: Mitsubishi Research Institute, 2007 (in Japanese).
- [2] Masui, T. "Framework of means for Environmental Improvement in Logistics." *Proceedings of the 9th International Symposium on Logistics*, Bangalore, India, July 2004, 504-510.
- [3] Masui and Goto Laboratory. *Environmental Loads Grasp and Logistics Strategies in Logistics Process*. Yokohama, Japan: Research Report, 2007 (in Japanese).
- [4] Masui and Goto Laboratory. *An Estimation of CO2 Emission in Logistics Process – A Case Study*-. Research Report, 2006 (in Japanese).
- [5] Suzuki, K., Masui, T., Goto, M., Nakahara, H. "A Survey on the Customer's Sense Towards Environmental Logistics and its Application to the Sales and Delivery System." *Proceedings of the 12th International Symposium on Logistics*, Budapest, Hungary, July 2007, 368-375.

Technology and the Purchasing Function & Cycle

Arun Madapusi
University of North Texas
Department of Management, College of Business Administration
P.O. Box 305429, Denton, Texas 76203.
Phone: (940) 565 3338; E-mail: madapusi@unt.edu

Abstract

The purpose of this research study is to examine the impact of technology on a firm's purchasing cycle. Three technology-induced paradigm shifts can be clearly discerned in the purchasing function and these are examined with respect to their impact on the purchasing function and cycle.

Keywords: Technology, Purchasing, Enterprise Systems.

Introduction

Purchasing has traditionally been a reactive function and considered a non-value-adding administrative task in most organizations. Firms typically followed a nine step process to address their purchasing needs – recognition of need, description of need, identification and analysis of sources of supply, supplier selection and determination of terms, preparation and placement of purchase order, follow-up and expediting of order, receipt and inspection of goods, invoice clearing and payment, maintenance of records and relationships [4] [5]. The introduction of technologies into the purchasing function, particularly in the past decade, has radically transformed the activities associated with each step of the purchasing cycle.

Impact of Paradigm Shifts on the Purchasing Function & Cycle

Three technology-induced paradigm shifts can be clearly discerned in the purchasing function with its concomitant impact on the purchasing cycle. The first paradigm shift was a tactical one and dealt with achieving efficiencies within the firm through the deployment of enterprise systems. Materials requirements planning (MRP) systems addressed the single need for materials planning in organizations. Manufacturing resource planning (MRP II) systems extended the scope of MRP systems to encompass the entire manufacturing function. MRP II systems then gave way to enterprise resource planning (ERP) systems which integrated all of a firm's functional activities [6].

The implementation of these enterprise systems enabled firms to streamline purchasing activities associated with each step of the purchasing cycle. Firms, however, found that competing in the global marketplace called for integrated planning and balancing of supply and demand across the supply chain. This gave rise to a second paradigm shift wherein the focus of the purchasing function shifted from a tactical role to a strategic one with accompanying impacts on the purchasing cycle. Firms sought to extend their ERP systems across the supply chain with the deployment of modules that catered to inter-firm activities. These extended ERP systems resulted in the creation of a seamless supply chain with the major thrust being supply management - building supplier partnerships, reduction of the supplier base, integration of supplier scheduling techniques with just-in-time, supplier certification, supplier quality assurance, and performance measurement [1] [3].

The extended ERP systems, though effective in managing the production purchasing needs of organizations, by and large ignored the firms' non-production purchasing needs such as maintenance, repair, and office (MRO) supplies. Firms typically kept the purchasing activities associated with MRO items outside the purview of the purchasing cycle. To address this lacuna and bring the MRO purchasing process within the gambit of the purchasing cycle, firms deployed e-procurement systems. This third paradigm shift, involving the use of web-enabled procurement systems and their integration with enterprise systems, enabled firms to realize reduced prices for materials and services, shortened acquisition and fulfillment times, decreased costs, improved inventory practices, and increased control over off-contract purchases [2].

Each of the three paradigm shifts differentially as well as collectively impact the various steps in the purchasing cycle. For example, in the first two steps of the purchasing cycle – recognition and description of need – the individual heads of departments receive standard purchase requisitions from work-units that are typically automatically generated through a re-order point system using an enterprise system or through an e-procurement system from individual employees in the case of MRO purchases. As another example, in step five – preparation and placement of purchase order – the purchasing function transmits the order to the supplier by electronic data interchange (embedded in the enterprise systems).

The three paradigm shifts discussed above not only streamlined purchasing activities at each step of the purchasing cycle but also compressed the cycle by speeding up and eliminating entire steps. As firms experiment with new technologies – supply portals, product exchanges, supply chain exchanges, web-based self-service, mobile-commerce – and extend their supply chains across countries, there are indications that a fourth paradigm shift is under way that could radically transform the purchasing cycle and enable firms to position their purchasing function as a key source of competitive advantage.

References

- [1] Akkermans, H.A., Bogerd, P., Yucesan, E. and Van Wassenhove, L.N. "The Impact of ERP on Supply Chain Management: Exploratory Findings from and European Delphi Study." *European Journal of Operational Research*, 2003, 146 (2), 284-301.
- [2] Ash, C.G., Burn, J.M. "A Strategic Framework for the Management of ERP-enabled e-Business Change." *European Journal of Operational Research*, 2003, 146 (2), 374-387.
- [3] Ayers, J. "Is Supply Management the Same as ERP?" *Information Strategy: The Executive's Journal*, 2001, 17(3), 43-45.
- [4] Burt, D.N., Dobler, D.W., Starling, S.L. *World Class Supply Chain Management*. 7th. Ed., NY: McGraw-Hill, 2003.
- [5] Leenders, M.R., Johnson, F.P., Flynn, A, E., Fearon, H.E. *Purchasing and Supply Management*. 13th. Ed., NY: McGraw-Hill, 2006.
- [6] Rondeau, P.J., Litteral, L.A. Evolution of Manufacturing Planning and Control Systems: From Reorder Point to Enterprise Resource Planning. *Production and Inventory Management Journal*, 2001, 42 (2), 1-7.

LOGARITHMIC INVESTMENT TO LOWER THE PROBABILITY OF SHORTAGE IN A CONTINUOUS-REVIEW (s,Q) MODEL WITH EXPONENTIAL DEMAND

Farrokh Nasri, Hofstra University, (516)463-4244, Farrokh.Nasri@hofstra.edu

Javad M. Paknejad, Hofstra University, (516)463-4244, M.J.Paknejad@hofstra.edu

John F. Affisco, Hofstra University, (516)463-4244, John.F.Affisco@hofstra.edu

ABSTRACT

One of the main challenges faced by the managers of inventory systems is when unsatisfied demand or shortage is realized as a result of demand exceeding the available stock. Inventory shortages may be determined by management of these systems to meet a given service level. Shortages result into significant opportunity costs, and even perhaps actual costs. These costs depend upon the structure of the mathematical model describing the behavior of the inventory system. In a recent paper, Paknejad, Nasri, and Affisco [4] provided the management the option of investing to lower the probability of shortage, and thus reducing the uncertainty on the availability of supply, improving the service level, and lower the relevant cost. Specifically, the authors considered a well known continuous-review (s,Q) model and investigated the impact of investing to reduce the upper bound of shortage probability based on the assumptions that (1) distribution of demand during the lead-time is uniform over the range 0 to t and (2) the probability of shortage follows an inverse investment function. The main objective of this paper is to extend the results in [4] when the density function of demand during the lead-time is exponential and the probability of shortage follows a logarithmic investment function. Explicit results for the optimal values of policy variables as well as the optimal value of the probability of shortage are obtained. A numerical example is also given.

INTRODUCTION

Managers of inventory systems are primarily paid to handle mismatches between supply and demand processes that do not fit neatly together. The most challenging of these mismatches is when unsatisfied demand or shortage is realized as a result of demand exceeding the available stock. Inventories are often held, then, to prevent or reduce shortages. These inventories may be in the form of safety stock or safety lead-time. In either case, a higher chance or probability of shortage requires a larger buffer; that is, larger inventories in either material or time. The question of shortages can be considered a matter of policy which is directly related to the service level at which management decides to operate. A typical policy, which is adopted in a recent

paper by Paknejad, Nasri, Affisco [4], is to constrain the probability of shortage by subjecting it to an admissible upper bound. The selection of the specific value for the upper bound is an indication of the trade-off between service level and cost that management is willing to accept. The value selected must reflect the realities of quality problems, lead-time variability, errors in shipments, transportation delays, forecasting errors, scheduling problems, inventory tracking problems, etc.

Ideally, management would like to be able to set the upper bound on the probability of shortage to an extremely small value, thus reducing the uncertainty of the availability of supply and improving the service level. One way to accomplish this is invest in programs aimed at correcting some of the problems mentioned above. In a recent paper, Paknejad, Nasri, and Affisco [4] presented a mathematical model to investigate the impact of such investment. Specifically, the authors considered a well known continuous-review (s,Q) model and investigated the impact of investing to reduce the upper bound of shortage probability based on the assumptions that (1) distribution of demand during the lead-time is uniform over the range 0 to t and (2) the probability of shortage follows an inverse investment function. The main objective of this paper is to extend the results in [4] when the density function of demand during the lead-time is exponential and the probability of shortage follows a logarithmic investment function.

BASIC MODEL AND PRELIMINERIES

The basic model considered in Paknejad, Nasri, and Affisco [4] is the continuous-review (s,Q) model studied by many authors including Hadley and Whitin [1], Johnson and Montgomery [3], Wagner [5], and Hillier and Lieberman [2]. In this model, Q units are ordered whenever the

inventory position (the amount on-hand plus the amount on-order minus the amount backordered) reaches the reorder level, s . It is assumed that demand for the item is stochastic and there is a fixed delivery time of known length before the order is received. Assuming that demand during lead-time is a continuous random variable, having a specific probability density function and using notations similar to those in Hillier and Liberman [2], the expected total cost per unit time, $c(Q,s)$, is shown to be

$$c(Q,s) = \frac{DK}{Q} + h \left[\frac{Q}{2} + s - D\lambda \right] + \frac{\pi D}{Q} \int_s^{\infty} (x-s) f(x) dx, \quad (1)$$

where

Q = order quantity per order,

s = reorder level,

D = expected number of units demanded per unit time,

λ = lead-time, time between placement and receipt of the order,

x = demand during the lead-time, a continuous random variable,

$f(x)$ = probability density function of demand during the lead-time,

$\mu = D\lambda$ = expected demand during lead-time,

K = ordering cost per order,

h = holding cost per unit per unit time, and

π = shortage cost per unit short, independent of the duration of the shortage.

The optimum values of the decision variables Q^* and s^* are found by setting the partial derivatives of (1) with respect to Q , and s equal to zero and solving as follows

$$Q^* = \sqrt{\frac{2D \left[K + \pi \int_{s^*}^{\infty} (x-s^*) f(x) dx \right]}{h}} \quad (2)$$

and

$$\int_{s^*}^{\infty} f(x) dx = \frac{hQ^*}{\pi D} \quad (3)$$

Since there is no closed-form solution for equations (2) and (3), Paknejad, Nasri, and Affisco [4] considered a uniform distribution of demand during lead-time and sought to choose an optimal policy subject to the constraint that the probability of a shortage occurring during a cycle will be less than or equal to an upper limit, V , set by the management. They then considered the option of investing in reducing the V parameter in the basic continuous-review (s, Q) inventory model when demand during lead-time is uniform in the range 0 to t and the probability of shortage follows an inverse investment function. Specifically, they found the following results for optimal probability of shortage, optimal production lot size, optimal reorder level, and the optimal expected cost:

$$V^{**} = \min \{V_0, V^*\} \quad (4)$$

$$Q^{**} = \min \{Q_0, Q^*\} \quad (5)$$

$$s^{**} = \max \{s_0, s^*\} \quad (6)$$

where

V_0 = Original upper bound of the probability of shortage

$$V^* = \sqrt{\frac{2 \left[\frac{Kh}{\pi} + ia \right]}{\pi D - ht}} \quad (7)$$

$$Q_0 = \sqrt{\frac{2DK + D\pi t V_0^2}{h}} \quad (8)$$

$$Q^* = \sqrt{\frac{2D}{h} \left[K + t \left(\frac{Kh + ia\pi}{\pi D - ht} \right) \right]} \quad (9)$$

$$s_0 = t(1 - v_0) \quad (10)$$

$$s^* = t \left(1 - \sqrt{\frac{2 \left[\frac{Kh}{\pi} \right] + ia}{\pi D - ht}} \right) \quad (11)$$

$$c^{**} = \min \{ c_0, c^* \} \quad (12)$$

where

$$c_0 = \frac{Kh}{\pi V_0} + \frac{\pi D V_0}{2} + h \left(t - \frac{t V_0}{2} - D \lambda \right) \quad (13)$$

and

$$c^* = i \left(\frac{a}{V^*} - b \right) + \frac{Kh}{\pi V^*} + \frac{\pi D V^*}{2} + h \left(t - \frac{t V^*}{2} - D \right) \quad (14)$$

In this paper we extend the results in [4] by considering the more realistic case of exponential probability density function for demand during the lead-time assuming that the probability of shortage follows a logarithmic investment function. It should be noted that the logarithmic investment function is extensively used in previous research by many authors, including the authors of this paper, in models dealing with quality improvement, lead-time variability reduction, and setup cost reduction. Its use in this study, in the context of probability of shortage reduction, is justified based on the idea that efforts aimed at lowering the probability of shortage most likely exhibit decreasing marginal return.

CASE OF EXPONENTIAL DEMAND DURING THE LEAD-TIME

If the density function of demand during lead-time is exponential, i.e.

$$f(x) = \frac{1}{D\lambda} e^{-\frac{x}{D\lambda}} \quad \text{for } x > 0 \quad (15)$$

then

$$\int_s^{\infty} f(x) dx = e^{-\frac{s}{D\lambda}} \quad (16)$$

and

$$\int_s^{\infty} (x-s) f(x) dx = D\lambda e^{-\frac{s}{D\lambda}} \quad (17)$$

Using the above relationships and minimizing $C(Q,s)$ of equation (1) subject to the constraint that $p(x>s) \leq V$, one can easily show that the optimal values of order quantity, reorder level, and expected total cost are

$$Q_o = \sqrt{\frac{2DK}{h} + \frac{2\pi D^2 \lambda V}{h}} \quad (18)$$

$$s_o = -D\lambda \ln V \quad (19)$$

and

$$c(Q_o, s_o) = \sqrt{2Dh[K + \pi D\lambda V]} - D\lambda h(1 + \ln V) \quad (20)$$

It can easily be shown that $C(Q_o, s_o)$ is a strictly convex function of V when $V > h\lambda/\pi$. This function reaches its minimum at

$$V = \frac{h\lambda + \sqrt{h^2 \lambda^2 + \frac{2hK}{D}}}{\pi} \quad (21)$$

The V in equations (18) through (20) is a parameter which represents the upper limit of the probability of shortage occurring during a cycle. We assumed that this parameter is constant. However, as noted in [4], it does not have to be.

THE OPTIMAL UPPER LIMIT OF PROBABILITY OF SHORTAGE

We now consider the option of investing to reduce the V parameter in the model of previous section. Following the approach in [4], we treat V as a policy variable and pursue the objective of minimizing the sum of the investment cost of changing V and the inventory related costs.

Specifically, we seek to minimize

$$C(Q, s, V) = ia_v(V) + c(Q, s) \quad (22)$$

subject to

$$p(x > s) \leq V_0 \quad (23)$$

where i is the cost of capital, $a_v(V)$ is a convex and strictly decreasing function of V representing the cost of changing the upper value of the probability of shortage to the level V , $c(Q, s)$ is the total inventory cost given by equation (1), and V_0 is the original V value before any investment is made to reduce it.

To solve this constrained optimization problem, we use a rather unorthodox approach of fixing V , optimizing over Q and s to obtain $c(Q_0, s_0)$ given by equation (20), and then minimizing over V . In this case our objective is to minimize

$$C(V) = ia_v(V) + c(Q_0, s_0) \quad (24)$$

In what follows we assume that $a_v(V)$ is a logarithmic investment function of V given by

$$a_v(V) = a - b \cdot \ln(V) \quad \text{for } 0 < V \leq V_0 \quad (25)$$

where

$$a = \ln V_0 / \Delta \quad (26)$$

and

$$b = 1/\Delta \quad (27)$$

and Δ is the percentage decrease in V per dollar increase in a_v .

THEOREM: If Δ is strictly positive, then the following hold

- a) $C(V)$ is strictly convex in V for all values of V greater than $h\lambda/\pi$.
- b) The optimal probability of shortage, optimal production lot size, optimal reorder level,

and the optimal expected cost are as follows:

$$V^* = \min \{V_o, V^*\}$$

$$Q^* = \min \{Q_o, Q^*\}$$

$$s^* = \min \{s_o, s^*\}$$

$$C^* = \min \{C_o, C^*\}$$

where

V_o = original upper bound of probability of shortage

$$V^* = \frac{1}{\pi} \left(1 + \frac{ib}{Dh\lambda} \right) \left[\left(h\lambda + \frac{ib}{D} \right) + \sqrt{\left(h\lambda + \frac{ib}{D} \right)^2 + \frac{2Kh}{D}} \right] \quad (28)$$

$$Q^* = \sqrt{\frac{2DK}{h} + \frac{2\pi D^2 \lambda}{h} V^*} \quad (29)$$

$$s^* = -D\lambda \ln V^* \quad (30)$$

$$C^* = b \ln \frac{V_o}{V^*} + \sqrt{2Dh[K + \pi D\lambda V^*]} - D\lambda h(1 + \ln V^*) \quad (31)$$

and Q_o, s_o and C_o are given by equations (18), (19), and (20) with $V=V_o$, respectively.

Proof is omitted.

NUMERICAL EXAMPLE

Suppose that demand for an item is exponentially distributed with a mean of 6,000 units per year.

Further assume that the delivery lead-time, λ , is one month, and the following parameters are

known: $K=\$1,000$, $h=\$2$ per unit per year, $\pi=\$100$ per unit short, and $\mu=D\lambda=6,000(1/12)=500$

units per month. Assuming that the management has an initial upper bound of 0.20 for the probability of shortage, V , the problem is whether or not to invest in efforts to reduce V if the cost of capital is 10 percent and the investment function follows equation (25) with $\Delta=0.01$.

In this case, when we do not have the option of investing to reduce V (i.e, $V=0.20$), the resulting optimal lot size is 8124, the optimal reorder point is 805, and the optimal expected total cost is \$16,858. Using the results of the theorem of the paper, the optimal value of V becomes 0.011 which yields an optimal lot size of 3,060, an optimal reorder point of 2,245, and an optimal expected total cost of \$12,491 per year. Thus there is a reduction in the expected total cost of \$4,367 per year. Reduction in V from the initial value of 0.20 to the optimal value of 0.011 requires an investment of \$28,800 which costs \$2,880 per year. The total return (before the investment cost) amounts to \$7,246 per year.

CONCLUSION

This paper presents an extension of a recent paper by Paknejad, Nasri, and Affisco [4] which considered a well known continuous-review (s,Q) model and investigated the impact of investment to reduce the upper bound of shortage probability for uniform demand during lead-time and an inverse investment function. Specifically, this paper extends the results in [4] to the case of exponential demand during lead-time and logarithmic investment cost function for reducing the probability of shortage. The paper presents explicit results for the optimal values of decision variables and the optimal expected annual cost.

References will be submitted upon request.

β = the fraction of stockouts that will be backordered

Variables

Q = the order quantity

T = the length of an order cycle

I = the maximum inventory level, with \bar{I} being the average inventory level over the year

S = the maximum stockout level, including both backorders and lost sales

B = the maximum backorder position, with \bar{B} being the average backorder level over the year ($B = \beta S$)

F = the fill rate or the percentage of demand that will be filled from stock

THE “PURE” STOCKOUT CASES: BACKORDERS AND LOST SALES

The EPQ model under the assumption that all stockouts are backordered at a cost C_b per unit per year appears in many basic texts. The optimal values for Q^* , the order quantity, B^* , the maximum backorder quantity, and T^* , the time between orders, are:

$$Q^* = \sqrt{\frac{2C_o D}{C_h(1-D/P)}} \sqrt{\frac{C_b + C_h}{C_b}}, B^* = Q^* \left(1 - \frac{D}{P}\right) \left(\frac{C_h}{C_b + C_h}\right), T^* = \sqrt{\frac{2C_o}{DC_h(1-D/P)}} \sqrt{\frac{C_b + C_h}{C_b}} \quad (1)$$

Q^* and T^* are those given by the basic EPQ formula, inflated to reflect the relative sizes of the unit inventory cost per year and the unit backorder cost per year and B^* is a fraction of Q^* that depends on the relative sizes of those two costs and the relative production and demand rates.

It is possible to prove for the basic EOQ that if demands occurring during a stockout period are lost sales rather than backorders, the optimal policy is to either have no stockouts or all stockouts, depending on which costs less. The same result can be proved for the EPQ.

MODELS FOR THE EOQ AND EPQ WITH PARTIAL BACKORDERING

Since, for the EOQ, it is optimal to allow some stockouts if all customers will wait ($\beta = 1$) and it is optimal to either allow no stockouts or lose all sales if no customers will wait ($\beta = 0$), it is logical that there will be a value of β below which one should use the optimal ordering policy for the lost-sales case – either using the basic EOQ model or never ordering at all, depending on which alternative is less costly – and above which one should allow stockouts, some of which will be backordered. Determining an optimal policy for the partial backordering EOQ problem starts with determining the minimum value of β for which stockouts should be allowed and, if β is greater than this minimum value, determining the optimal order quantity.

Models for the partial backordering EOQ problem were developed by Montgomery et al. [2], Rosenberg [6], Park [3], San José et al. [7], and Pentico and Drake [4]. These papers took somewhat different approaches to modeling the problem, differing primarily in which decision variables they focused on, although San José et al. [7], rather than assuming that β is a constant, considered a number of different “customer impatience” functions that have the property that the percentage backordered does not decrease as the replenishment date approaches. The first four papers resulted in complicated solution procedures. Pentico and Drake [4] took a different approach that resulted in a model that is more comparable to the EOQ with full backordering.

Mak [1] and Pentico and Drake [5] developed models for the EPQ with partial backordering. Other than notation, they differ in their policies on filling demands that occur while there is no stock but the production run has started, which has implications for modeling the problem.

Modeling Inventory For The EPQ With Backordering

With partial backordering, from the time the system runs out of stock until the time the next order is received (EOQ) or the next production run begins (EPQ), a fraction β of incoming demand will be backordered until the maximum backorder level $B = \beta S$ is reached.

In the EOQ models with full or partial backordering, the entire order quantity Q is received at the same time, so all the backorders can be filled at once, with the maximum inventory rising immediately to $I = Q - B$. In the EPQ model with full backordering, the order quantity Q is received in a constant stream at a rate of P . Since all demands that occur during the time it takes to fill all the backorders are also backordered if they are not filled immediately, it makes no difference whether the incoming orders are filled before the backorders (LIFO) or the backorders are filled before the incoming orders (FIFO). Inventory increases at the rate of $P - D$ until the backorder is eliminated and the maximum inventory level, $I = Q(1 - P/D) - B$, is reached.

For the EPQ with partial backordering, however, whether LIFO or FIFO is used to determine the order in which demands are filled after the production run begins can make a difference in the net inventory level. Whether it does or not depends on the answer to an additional question: What happens to the demands that occur when there is no stock on hand but the production run has been started? If one assumes that incoming demands will be filled before the existing backorders (LIFO) and further assume that none of the existing backorders will convert to lost sales, then the net inventory level for the EPQ with partial backordering will increase at a rate of $P - D$. This is Mak's [1] unstated assumption. If one assumes, as Pentico and Drake [5] did, that the existing backorders will be filled before any new demands (FIFO) and further assume that only a fraction β of these new orders that cannot immediately be filled will be backordered, with the rest being lost sales, then the net inventory level will increase at a rate of $P - \beta D$ until the backorder is eliminated and then will increase at a rate of $P - D$ until the maximum inventory level is reached. (If all incoming orders will wait once the production run has started, it makes no difference whether LIFO or FIFO is used.)

Mak's [1] model for the EPQ with partial backordering

Mak's assumptions are the usual ones for the EPQ model with full backordering except that only a fraction β of the stockouts will be backordered, with the rest being lost sales. As noted above, he implicitly assumes that there will be no increase in either backorders or lost sales once the production phase begins, so the backorders are filled at a rate of $P - D$.

Mak's decision variables are T , the length of an inventory cycle, and t , the length of time from when the inventory level reaches 0 until the next production run begins (which is t_1 in the analysis to follow). His cost function is convex, so the optimal solution can be found by setting the two partial derivatives equal to 0 and solving the resulting equations simultaneously. He develops an equation for T as a function of t and, using this to eliminate T from one of the equations, finds an expression for t^* as a function of the parameters, which he uses to find an

expression for T^* . Both equations are complicated. Mak's statement of a condition that β must satisfy for the partial backordering EPQ equations to apply is not as simple as the ones developed for the partial backordering EOQ models or the one developed by Pentico and Drake [5].

Pentico and Drake's [5] model for the EPQ with Partial Backordering

In contrast with Mak [1], Pentico and Drake assumed a FIFO policy on filling backorders. Using T and F as their decision variables, they developed the following equations for an optimal policy:

$$T^* = \sqrt{\frac{2C_o}{DC'_h} \left[\frac{C'_h + \beta C'_b}{\beta C'_b} \right] - \frac{[(1-\beta)C_1]^2}{\beta C'_h C'_b}} \quad \text{and} \quad F(T) = \frac{(1-\beta)C_1 + \beta C'_b T}{T(C'_h + \beta C'_b)} \tag{2}$$

where $C'_h = C_h(1 - D/P)$ and $C'_b = C_b(1 - \beta D/P)$. They also developed a statement of the condition that β must meet for partial backordering to be optimal which, with the replacement of C_h by C'_h , is the same as in Pentico and Drake [4] for the EOQ with partial backordering:

$$\beta \geq \beta^* = 1 - \sqrt{\frac{2C_o C'_h}{DC_1^2}} \tag{3}$$

AN ALTERNATIVE MODEL FOR THE EPQ WITH PARTIAL BACKORDERING AND A LIFO POLCIY

We develop here an alternative to Mak's [1] model for the EPQ with partial backordering and a LIFO policy on filling backorders. We will use the same modeling approach used by Pentico and Drake [4,5] for the EOQ with partial backordering and the EPQ with partial backordering and a FIFO policy on filling backorders. We will show how the use of the LIFO rather than the FIFO policy changes the equations and the condition β must meet for those equations to give an optimal solution.

The Profit and Cost Functions Based on T and F

The first step is to develop equations for the lengths of four parts of the inventory cycle. Starting with the time that the existing inventory is exhausted, they are: 1) t_1 is the time until the maximum backorder level is reached and production starts, 2) t_2 is the time from the start of production until the backorder is eliminated and inventory starts to accumulate, 3) t_3 is the time until production stops and the maximum inventory level I is reached, 4) t_4 is the time until the inventory is exhausted and a new cycle begins. Under LIFO the values for $t_1, t_2, t_3,$ and t_4 are:

$$t_1 = \frac{(1-F)T(P-D)}{P-D(1-\beta)}, \quad t_2 = \frac{\beta(1-F)TD}{P-D(1-\beta)}, \quad t_3 = \frac{FTD}{P}, \quad t_4 = FT\left(1 - \frac{D}{P}\right) \tag{4}$$

The average profit per year to be maximized is the revenue from filling demands, either from stock or as backorders, minus the cost of placing orders, the cost of the units used or sold, the cost of carrying inventory, the cost of the backorders, and the cost of lost sales. This is maximized by the pair (T,F) that minimizes the average cost per period:

$$\Gamma(T,F) = \frac{C_o}{T} + C_h \bar{I} + C_b \bar{B} + C_1 D(1-\beta)(1-F) \tag{5}$$

Since $I = FTD(1 - D/P)$ and $B = \beta Dt_1 = \beta D(1 - F)T(P - D)/(P - D(1 - \beta))$, the average inventory and average backorder level are:

$$\bar{I} = \frac{DTF^2}{2} \left(1 - \frac{D}{P}\right) \text{ and } \bar{B} = \frac{\beta(1-F)^2TD(P-D)}{2(P-D(1-\beta))}. \tag{6}$$

Substituting the expressions for \bar{I} and \bar{B} into (5) gives:

$$\Gamma(T,F) = \frac{C_o}{T} + \frac{C_hDTF^2}{2} \left(1 - \frac{D}{P}\right) + \frac{\beta C_bDT(1-F)^2(P-D)}{2(P-D(1-\beta))} + \frac{C_lD(1-F)(1-\beta)(P-D)}{P-D(1-\beta)}. \tag{7}$$

Determining the optimal values for T and F

Taking the partial derivative of $\Gamma(T,F)$ in (7) with respect to T and setting it equal to 0 gives:

$$T = \sqrt{\frac{2C_o}{D[F^2C_h(1-\frac{D}{P})(1-\frac{D(1-\beta)}{P}) + (1-F)^2\beta C_b(1-\frac{D}{P})]}}. \tag{8}$$

Taking the partial derivative of $\Gamma(T,F)$ with respect to F and setting it equal to 0 gives:

$$F = \frac{(1-\beta)C_l + \beta C_bT}{T(C_h(1-\frac{D(1-\beta)}{P}) + \beta C_b)}. \tag{9}$$

We define $C_o^* = C_o(1 - D(1 - \beta)/P)$ and $C_h^* = C_h(1 - D(1 - \beta)/P)$. With these replacements for C_o and C_h , substituting the expression for F in (9) into equation (8), we get, after some algebra:

$$T^* = \sqrt{\frac{2C_o^*}{DC_h^*(1-D/P)} \left[\frac{C_h^* + \beta C_b}{\beta C_b} \right] - \frac{[(1-\beta)C_l]^2}{\beta C_h^* C_b}}. \tag{10}$$

$$F(T^*) = \frac{(1-\beta)C_l + \beta C_b T^*}{T^*(C_h^* + \beta C_b)}. \tag{11}$$

Determining the minimum value of β for optimality

Recognizing that T^* for the partial backordering model must be at least as large as T^* for the basic EPQ $\left(\sqrt{2C_o/(DC_h(1-D/P))}\right)$ for backordering to be optimal gives the bound:

$$\frac{2C_o^*}{DC_h^*(1-D/P)} \left[\frac{C_h^* + \beta C_b}{\beta C_b} \right] - \frac{[(1-\beta)C_l]^2}{\beta C_h^* C_b} \geq \frac{2C_o}{DC_h(1-D/P)}.$$

After some algebra, this leads us to the following conclusion: For the equations for T^* and F^* to give an optimal solution, we must have:

$$\sqrt{\frac{2C_o}{DC_h(1-D/P)}} \geq \frac{(1-\beta)C_l}{C_h^*} \text{ or } \beta \geq 1 - \sqrt{\frac{2C_o^* C_h^*}{DC_l^2(1-D/P)}}. \tag{12}$$

This is the same criterion derived by Mak [1]. We note that it has the same basic form as the condition derived by Park [3] and Pentico and Drake [4] for determining the minimum value of β for which backordering is optimal for the partial backordering EOQ.

Procedure for finding the optimal inventory policy

The procedure for determining the optimal values for T, F, Q, I, S, and B is, then:

1. Use (12) to determine whether β meets the criterion for the optimality of partial backordering.

2. a. If “No”, determine T^* from the basic EPQ model ($T^* = \sqrt{\frac{2C_o}{C_h D(1-D/P)}}$) and determine the optimal cost of allowing no stockouts ($\Gamma^* = \sqrt{2C_o C_h D(1-D/P)}$). Compare this with the cost of losing all demand, $C_1 D$, to determine whether it is optimal to allow no stockouts or all stockouts.
- b. If “Yes”, use (10) to determine the value of T^* and substitute it into (11) to determine the value of F^* .
3. Determine the values of the other variables as follows:
 - Total demand during a cycle = DT^*
 - Maximum inventory = $I^* = F^*DT^*(1 - D/P)$
 - Maximum stockout = $S^* = Dt_1^* = (1 - F^*)T^*D(P - D)/(P - D(1 - \beta))$
 - Maximum backorder = $B^* = \beta S^*$
 - Order quantity = $Q^* = DT^* - (1 - \beta)S^*$

CONCLUSION

By using T , the time between orders, and F , the fill rate, as the decision variables, we have developed a model for the EPQ with partial backordering and a LIFO policy on filling backorders with equations that are more like those for the basic EPQ model and its full-backordering extension and are easier to solve than the equations developed by Mak [1].

REFERENCES

- [1] Mak, K. L. “Determining optimal production-inventory control policies for an inventory system with partial backlogging.” *Computers & Operations Research*, 1987, 14(4), 299-304.
- [2] Montgomery, D.C., Bazaraa, M.S. and Keswani, A.K. “Inventory models with a mixture of backorders and lost sales.” *Naval Research Logistics Quarterly*, 1973, 20(2), 255-263.
- [3] Park, K. S. “Inventory model with partial backorders.” *International Journal of Systems Science*, 1982, 13(12), 1313-1317.
- [4] Pentico, D.W. and Drake, M.J. “The deterministic EOQ with partial backordering: a new approach.” *European Journal of Operational Research*, 2007 (in press).
- [5] Pentico, D.W. and Drake, M.J. “The deterministic EPQ with partial backordering: a new approach.” 2007, *Proceedings*, Northeast DSI Conference, Baltimore, 521-526.
- [6] Rosenberg, D. “A new analysis of a lot-size model with partial backordering.” *Naval Research Logistics Quarterly*, 1979, 26(2), 349-353.
- [7] San José, L.A., Sicilia, J. and García-Laguna, J. “The lot size-reorder level inventory system with customers impatience functions.” *Computers & Industrial Engineering*, 2005, 49(3), 349-362.

THE EPQ WITH PARTIAL BACKORDERING AND A LIFO POLICY ON FILLING BACKORDERS

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

ABSTRACT

The first published model for the EPQ when only a percentage of stockouts will be backordered was by Mak [1]. He assumed that, once production started, new customer orders would be filled before the existing backorders. The resulting equation set was complicated and non-intuitive. In this paper we use the approach taken by Pentico and Drake [4] for the EOQ with partial backordering to develop equations for the EPQ with partial backordering and a LIFO policy on filling backorders that are more like those for the EPQ with full backordering.

INTRODUCTION

A key assumption of the basic EOQ and EPQ models is that stockouts are not permitted. However, if customers are always willing to wait for delivery, planned backorders can make economic sense, even if they incur a cost. Relaxing the basic models' assumption that stockouts are not permitted led to the development of models for the two basic stockout cases: backorders and lost sales. What took longer to develop was a model for the case where only a percentage of customers are willing to wait for delivery and the rest will cancel their orders unless the supplier fills them within the normal delivery time by using more expensive supply methods.

After reviewing the models for the "pure" stockout cases of backorders and lost sales, we will briefly discuss the EPQ models with partial backordering by Mak [1] and Pentico and Drake [5]. We will present an alternative to Mak's approach to modeling the EPQ with a LIFO policy on filling backorders based on the approach in Pentico and Drake [4,5].

NOTATION AND TERMINOLOGY

Parameters

D = demand per year

P = production rate per year if constantly producing

s = the unit selling price

C_o = the fixed cost of placing and receiving an order

C_p = the variable cost of a purchasing or producing a unit

C_h = the cost to hold a unit in inventory for a year

C_b = the cost to keep a unit backordered for a year

C_g = the goodwill loss on a unit of unfilled demand

$C_1 = (s - C_p) + C_g$ = the cost for a lost sale, including the lost profit on that unit and any goodwill loss

USING SIMULATION OPTIMIZATION FOR PARTS RECOVERY AND SPARE PARTS INVENTORY MANAGEMENT IN POST PRODUCT LIFE CYCLE

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

The parts recovered from discarded and end-of-life products can be used as a source of spare parts during post product life cycle. However, accurate determination of the final order quantity is complicated as it requires the prediction of spare parts demand for the post product life cycle. In this paper, the spare parts problem is solved using simulation. Simulation is chosen because of the stochastic and complex nature of the problem. A simulation model of the manufacturing system is first developed which is then integrated with a genetic algorithm to determine the optimal final order quantities for a number of critical spare parts.

INTRODUCTION

The success of companies in after sales service largely depend on the timely and cost effective supply of spare parts which are needed in the repair and maintenance of durable products. The acquisition of spare parts can be an extremely challenging task especially at the post product life cycle since the source of new spare parts is limited.

Placing a final order for spare parts at the end of product life cycle is a common approach to this problem. However, accurate determination of this final order quantity is a complicated issue since it requires the prediction of spare part demand for the post product life cycle. Prediction of demand is very difficult even when the failure rates are available and the number of installed bases is known, because a long term forecast over several years is needed and a number of intangible factors such as fashion and obsolescence must be considered. If the demand is overestimated during the post product cycle, excess parts have to be scrapped. Underestimation of demand results in opportunity cost and lost goodwill of the customers [1].

One possible solution to the above spare parts problem is that the recovered parts from the discarded and end-of-life (EOL) products can be used as a source of spare parts during post product life cycle because remanufacturing cost is much cheaper than cost of extra production [2]. However, the remanufacturing system is fraught with high degree of uncertainties in timing, quantity and quality of product returns. Thus a satisfactory decision model for the determination of final order quantities for spare parts during parts recovery and disassembly must have the ability of coping with many stochastic variables simultaneously.

CURRENT STATUS

In the literature, the most commonly used approaches to develop spare provisioning decision models are simulation and mathematical programming. Mathematical programming consists of the development of mathematical models based on techniques such as linear programming, dynamic programming and goal programming. Sherbrooke's [3] multi-echelon model for recoverable item control is the first application of mathematical programming in spare parts inventory management problem. After that study, many researchers studied various aspects of the spare parts management problem. The reader is referred to Kennedy et al.'s [4] paper for an overview of these studies. It is

noted that all these studies entail the use of simplified plants or systems models whose predictions may be of questionable realism and reliability.

The use of simulation in modeling the spare parts inventory management problem represents a popular alternative to mathematical programming since simulation has the ability of describing multivariate non-linear relationships which can be difficult to put in an explicit analytical form. However, simulation modeling is not an optimization technique. It is, therefore, necessary to integrate the simulation model with an optimization tool.

In recent years, metaheuristics such as genetic algorithms (GAs), simulated annealing, and tabu search have been extensively used along with simulation to enhance the efficiency of the search procedure. Among these guided search methods, simulation optimization via GAs is quite an active research area.

GAs are biologically inspired search procedures that have been used to solve different NP-hard problems [5], [6], [7]. Like other biologically inspired techniques (e.g. ant colony optimization, particle swarm optimization), they try to extract ideas from a natural system, in particular the natural evolution, in order to develop computational tools for solving engineering problems. They are used to search large, non-linear search spaces where expert knowledge is lacking or difficult to encode and where traditional optimization methods fall short.

There are successful applications of GA-based simulation optimization in scheduling, facility layout, assembly-line planning, supply-chain management, kanban systems, maintenance-policy selection, and spare parts inventory management.

METHODOLOGY

This study tries to determine the optimal order quantities for a number of critical spare parts. Reordering of spare parts during post product life cycle is not possible in the system considered. The only source of spare parts during this period is the parts harvested from the discarded and EOL products. Hence, the determination of the optimal order quantities for critical spare parts requires the simultaneous consideration of the demand for spare parts during post product life cycle and timing, quantity and quality of product returns.

In order to cope with these highly uncertain components of the problem, firstly, a detailed simulation model describing the manufacturing system with its spare parts inventory and remanufacturing related aspects is developed. Uncertainty in demand during post product life cycle, uncertain product returns and part recovery is included in this simulation model. Then, a GA is integrated with the simulation model for the optimization of final order quantities for critical spare parts.

While designing the GA, at first the final order quantities of critical spare parts is coded into chromosomes so as to perform the genetic operation. Each chromosome represents a possible configuration of final order quantities of critical spare parts. The GA process involves searching for the optimum final order quantities of critical spare parts. In our approach, the fitness of each chromosome is evaluated by the simulation model. According to the fitness results, the GA creates new alternative solutions. Thus, there is a two-way communication between the GA and the simulation model.

During this search process, a total annual cost (TAC) function is employed to evaluate the fitness of each alternative solution. This annual cost function involves the holding and shortage cost of spare parts and the costs associated with product recovery and remanufacturing.

Next, to further improve the performance of the GA developed, a set of experiments is performed to identify appropriate values for the GA parameters (i.e. the size of the population, the crossover probability, and the mutation probability).

CONCLUSION

Accurate determination of the final order quantity is a complicated issue because it requires the prediction of spare part demand for the post product life cycle. Moreover, the recovered parts from the discarded products can be used as a source of spare parts during post product life cycle. This stochastic and complex nature of the problem requires the use of simulation modeling. In this study, first, a simulation model of the manufacturing system is developed. Then a GA is integrated with the simulation model to determine the optimal final order quantities for a number of critical spare parts.

REFERENCES

- [1] Schröter, M. and Spengler, T., “Designing control management systems for parts recovery and spare parts management in the final phase within closed-loop supply chains”, *International Journal of Integrated Supply Management*, 1 (2), 158-179, 2004.
- [2] Inderfurth, K. and Mukherjee, K., “Analysis of spare part acquisition in post product life cycle”, *Working Paper No. 6*, February 2006.
- [3] Sherbrooke, C. C., “METRIC: a multi-echelon technique for recoverable item control”, *Operations Research*, 16, 122–141, 1968.
- [4] Kennedy, W. J., Patterson, W. J. and Fredendall, L. D., “An overview of recent literature on spare parts inventories”, *International Journal of Production Economics*, 76, 201-215, 2002.
- [5] Kongar, E. and Gupta, S. M., “Disassembly Sequencing using Genetic Algorithm”, *International Journal of Advanced Manufacturing Technology*, 30, 497-506, 2006.
- [6] McGovern, S. M. and Gupta, S. M., “A Balancing Method and Genetic Algorithm for Disassembly Line Balancing”, *European Journal of Operational Research*, 179, 692-708, 2007.
- [7] Imtavanich, P. and Gupta, S. M., “Generating Disassembly Sequences for Multiple Products Using Genetic Algorithm”, *Proceedings of the 2007 POMS-Dallas Meeting*, Dallas, Texas, May 4-May 7, 2007 (CD-ROM).

AN ANALYTICAL MODEL FOR OPTIMAL DISASSEMBLY

Ying Tang

Electrical and Computer Engineering
Rowan University, Glassboro, NJ 08028
Tel: 856-256-5339; Email: tang@rowan.edu

ABSTRACT

To ease in the disassembly procedure of discarded products, this paper describes an analytical model of a Disassembly Petri net (DPN) in conjunction with a Hybrid Bayesian network (HBN), for optimal disassembly planning. The DPN models all feasible disassembly steps along with the necessary resources for the operations, whereas the HBN is a belief network that describes knowledge pertaining to various facets of the entire disassembly process (e.g., the defect rate of a product or the skill of a worker). While the two networks are linked together, pertinent knowledge can flow between them, allowing newly available information to be incorporated into the extended DPN for better decision making.

Keywords: Petri nets, Hybrid Bayesian Network, Optimal Disassembly

I. INTRODUCTION

According to the U.S. Environmental Protection Agency, an estimated 30 to 40 million PCs will be ready for "end-of-life management" in each of the next few years, and more than 70 percent of them eventually end up in landfills in the United States [1]. Considering the high risk of leaking toxic into the ground through dumping of such e-waste and goodly amounts of valuable metals contained by it, many groups are attempting to disassemble obsolete products in order to reuse or recycle their components/materials.

Disassembly process planning is a well-studied optimization problem that seeks to determine the level of disassembly of a reclaimed product to its constituent components with maximum returns. Many heuristic approaches [2, 5, 6, 7] have been developed to deal with this topic. In general, those methods work on a detailed representation of all feasible disassembly options (i.e., a mathematical model of the disassembly) of a reclaimed unit, associate this model with sort of data pertaining to various facets of the disassembly process (e.g., disassembly cost, disposition cost, and revenue of a component to be reused, etc.), and then perform a search algorithm that makes a sequence of decisions, starting from the product unit, to select from a set of feasible options that take the product apart into its constituent subassemblies and/or components, a venue that optimizes the expected returns. The challenge arises because much of the information necessary for disassembly decision making is not available during the process initiation and can only be provided through the real-time observation of the process itself. For

instance, obsolete products exhibit a high level of uncertainty in their structure and condition due to uncontrollable element of human behavior involved in product usage. And the influence of human factors on disassembly is imperative and hard to estimate *a priori*. Taking this into consideration, this paper proposed an analytical model for uncertainty management in optimal disassembly. The rest of the paper is organized as follows: Section II presents the model and its important role in optimal disassembly. Section III discusses the challenges associated with the model, followed by the conclusion in Section IV.

II. DISASSEMBLY PROCESS MODELING

The objective of disassembly is to maximize expected EOL values with respect to ecological, environmental and economic issues [8]. Various factors can lead to improved profits, such as reducing the overall time the disassembly process takes, choosing the most cost-efficient worker for a disassembly task, or deciding whether to resell/recycle a subassembly in lieu of further disassembly. For this to be successful, we first adopt the disassembly Petri net (DPN) model to represent all feasible disassembly options for a reclaimed product. Although the DPN has full knowledge of the actions that can be performed and the order that the actions are performed in, it does not necessarily know which action is the most appropriate to take since the information necessary to make such a decision is not readily available and tends to have a high degree of uncertainty. To alleviate this concern, a hybrid Bayesian network is developed to ascertain the uncertain parameters, such as the condition of a unit being processed and the time needed for a disassembly task. The two networks are detailed in the following subsections.

II. A. Disassembly Petri Net

An extended DPN is defined as a nine-tuple [4]:

DPN = (P, T, I, O, M, λ , τ , π_{resale} , $\pi_{disposal}$, φ , θ) where:

- $Z=(P, T, I, O, M)$ is an acyclic Petri net such that:
 - (1) P is a nonempty set, where $P = \{p_1, p_2, \dots, p_n\} = W \cup Q$. The set of places $W = \{w_1, w_2, \dots, w_r\}$ represents operators, and the set of places $Q = \{q_1, q_2, \dots, q_s\}$ stands for a product, subassemblies or components. $r + s = n$. $\bullet p$ is the set of input transitions to the place p , and p^\bullet is the set of output transitions to the place p .
 - (i) There is a unique $p \in Q$ with no input transitions (i.e., $\bullet p = \emptyset$): This place is usually denoted by p_1 and named as root or a product place.
 - (ii) There is a set of places called leaves denoted by Q , each of which has no output transitions (i.e., $p \in Q, p^\bullet = \emptyset$).
 - (2) $T = \{t_1, t_2, \dots, t_m\}$ is a nonempty set of transitions representing disassembly operations. $P \cap T = \emptyset$. $\bullet t$ is the set of input places to the transition t , and t^\bullet is the set of output places to the transition t
- $\lambda: T \rightarrow \mathfrak{R}^+$ is the expected disassembly time associated with a transition. Its unit is the hour.

- $\tau: Q \rightarrow \mathfrak{R}$ is the expected monetary value obtained from the resale, reuse, recycling, or disposal of the disassembled unit represented by $p, p \in Q$.
- $\pi_{resale}: Q \rightarrow \mathfrak{R}^+$ is the resale/reuse value of the disassembled unit represented by $p, p \in Q$ with no flaw
- $\pi_{disposal}: Q \rightarrow \{0\} \cup \mathfrak{R}^-$ is the disposal cost of the disassembled unit represented by $p, p \in Q$
- $\varphi: Q \rightarrow [0, 1]$ is the discrete probability of defect of a component/subassembly.
- $\theta: W \rightarrow \mathfrak{R}^+$ is operator wage rate associated with a place in W .

Note that π_{resale} and $\pi_{disposal}$ are determined by the prevailing market condition. Their values, as well as that of θ are assumed known *a priori*. As described in the next section, $\forall q \in Q$ and $\forall w \in W$, there is a correlative vertex in the corresponding HBN. Thus, π_{resale} , $\pi_{disposal}$ and θ are available for the necessary calculation in the HBN. On the other hand, the values of λ , τ , and φ are decided through the HBN learning.

II. B. Hybrid Bayesian Network

A Hybrid Bayesian network (HBN), $G = (V, E)$, is a directed acyclic graph

- $V = \{v_1, v_2, \dots, v_n\}$ is a set of vertices corresponding to random variables in the domain $V = \Delta \cup \Gamma$. The vertices in set Δ represent qualitative variables, and those in set Γ , quantitative variables.
- $E = \{(v_i, v_j) | v_i, v_j \in V, i \neq j\}$ is a set of ordered pairs of vertices indicating a probability dependency between the parent node v_i and the child node v_j . The network encodes a joint probability distribution of the domain variables $\mathcal{P}(v_1, v_2, \dots, v_n)$, where ψ_i denotes the parents of node v_i .

$$\mathcal{P}(v_1, v_2, \dots, v_n) = \prod_i \mathcal{P}(v_i | \psi_i)$$

In our approach, a HBN can be directly derived from its corresponding DPN. Each of the components/subassemblies represented in the DPN must have a node corresponding to its probability of defect in the HBN. The defect rate of the parent subassemblies will also influence the defect rate of the child components/subassemblies. The workers from the DPN must be added to the HBN as continuous value nodes, representing the salary of the workers. The salary of the workers is equivalent to their skill level, which influences the defect rate of the product being disassembled. The transition nodes from the DPN are represented as transition times in the HBN, which are influenced by the defect rate of the subassembly being disassembled. Finally, nodes corresponding to the quantitative and qualitative characteristics of the individual components of the product will be added to the HBN. By examining these characteristics and their relationship with the defect rates of the components, we use known characteristics to help obtain the probability of defect in a component.

III. CHALLENGE STATEMENT

This section discusses several challenges with this model and possible solutions.

When a HBN described in the previous section is generated, the probabilities of individual vertices and their dependency/independency relationships must be defined. As opposed to the homogenous BNs, a HBN proves to be more complicated in the incorporation of evidence, since most of the simplistic propagation schemes cannot handle a mixed network, especially when the network has discrete nodes with continuous parents. One of the solutions is to model such a node $R \in \{0, 1\}$ with continuous (vector-valued) parents X as softmax node [4]:

$$\mathbb{P}(R = 1 | X = x) = \frac{1}{1 + e^{-\eta}}$$

where $\eta = b + w'x$, w is the normal vector of weights and b is the decision boundary.

Further, obtaining the parameters for the softmax nodes, which represent the discrete probability of defect of the components, is not straightforward. Usually, when estimating for the parameters b and w , there is a set of training data available which has a multitude of vectors x and their resulting probabilities p . The training data is used in conjunction with a curve fitting algorithm to estimate the parameters. In our case, there is one caveat which must be taken into consideration. Typically, a component has multiple sets of training data where the values of the vector x are known. However, we only have knowledge of whether or not the component was defective and not the component's probability of defect. Without the probability of defect, we cannot use a curve fitting algorithm to determine the parameters for the corresponding logistic function. Furthermore, the probabilities of defect cannot be set to 0 or 1 for the curve fitting algorithm, since it would result in values of negative or positive infinity for the parameters. Instead, an approximation of the appropriate training data must be made in order for the parameters to be estimated. The detailed solution is referred to [3].

IV. CONCLUSION

Uncertainty management in disassembly optimization is emerging as a predominant issue in the field of reverse logistics. This paper addresses the modeling challenge by taking advantage of the prowess of the DPN in conjunction with the statistical inference provided by the HBN. We shall seek to implement our methods in software suite and provide it as a tool for others to use when studying and practicing optimal disassembly planning. A further investigation on efficient inference and search algorithms is another part of the future research.

ACKNOWLEDGEMENTS

This work is supported in part by the Charles A. and Ann Morrow Lindberg Foundation and the Christian R. and Mary F. Lindback Foundation.

REFERENCES

- [1] Carroll, C., "High Tech Trash: Will your discarded TV end up in a ditch in Ghana?" *National Geographic*, Jan. 2008.
- [2] Gao, M., Zhou, M. C. and Tang, Y., "Intelligent Decision Making in Disassembly Process Based on Fuzzy Reasoning Petri Nets," *IEEE Trans. on Systems, Man, and Cybernetics: Part B*, 2004 34(5), p. 2029-2034.
- [3] Grochowski, D. and Tang, Y., "Parameter Estimation for Optimal Disassembly Planning," *Proceedings of the IEEE International Conf. on Systems, Man and Cybernetics*, Montreal, Quebec, Canada, Oct. 7-10, 2007, pp.2490-2495
- [4] Grochowski, D. and Tang, Y., "A learning strategy for disassembly process planning," *Proc. of the IEEE International Conf. on Networking, Sensing and Control*, London, United Kingdom, April 15-17, 2007, pp. 489-495.
- [5] Lambert, A. J. D., "Determining optimum disassembly sequences in electronic equipment," *Computers and Industrial Engineering*, 2002, 43(3), p. 553-575.
- [6] Mascle, C. and Balasoiu, B. A., "Algorithmic selection of a disassembly sequence of a component by a wave propagation method," *Robotics and Computer Integrated Manufacturing*, 2003, 19, p. 439-448
- [7] Tang, Y., Zhou, M. C. and Caudill, R., "An Integrated Approach to Disassembly Planning and Demanufacturing Operation," *IEEE Transaction on Robotics & Automation*, 2001, 17(6), p. 773-784.
- [8] Zussman, E., Reiter, B. S., Scharke, H., "Modeling and planning of disassembly processes", *Proc. of the IFIP WG5.3 Inter. Conf. on Life-Cycle Modeling for Innovative Products and Processes*, Nov./Dec., 1995, pp. 221-232

DESIGN ISSUES FOR INVERSE MANUFACTURING SYSTEMS IN JAPANESE CASES

Tetsuo Yamada

Department of Environmental and Information Studies, Musashi Institute of Technology
Yokohama, Kanagawa, Japan, tyamada@yc.musashi-tech.ac.jp, (+)81-45-910-2916

ABSTRACT

Recent manufacturers should construct and design inverse manufacturing systems consisting of assembly and disassembly systems for environmentally-conscious manufacturing and sustainable manufacturing. In Japan, inverse manufacturing and disassembly systems have already been constructed and are being operated to produce and recover various kinds of products in practice. This study focuses on inverse manufacturing and disassembly systems in actual Japanese cases, and it discusses their system design by considering repeated product life cycles among product generations and types. First, the method of this survey and the classifications are defined for inverse manufacturing and disassembly systems. Next, the systems in Japanese cases are introduced and classified in terms of product life cycle. Finally, the issues of system design are discussed and identified.

Keywords: Disassembly/Assembly Systems, Reuse and Recycling, Repeated Product Life Cycle, Sustainable Manufacturing

INTRODUCTION

Recent manufacturers should construct and design inverse manufacturing systems consisting of assembly and disassembly systems for environmentally-conscious manufacturing [2] and sustainable manufacturing [6]. In Japan, inverse manufacturing and disassembly systems have already been constructed and are being operated to produce and recover various kinds of products such as recycling cameras, copiers, computers in practice. It is probably believed that inverse manufacturing systems can be constructed and realized for any kind of products and situations by reuse and recycling. However, it is impossible for the inverse manufacturing systems to be constructed without considering repeated product life cycles among product generations and types. This study focuses on inverse manufacturing and disassembly systems in actual Japanese cases, and it discusses their system design by considering repeated product life cycles among product generations and types. First, the method of this survey and the classifications are defined for inverse manufacturing and disassembly systems. Next, the systems in Japanese cases visited by the author are introduced and classified in terms of product life cycle. Finally, issues of system design are discussed and identified.

METHOD OF SURVEY

Recovery Circumstances

In Japan, there are already many inverse manufacturing and disassembly systems for various kinds of products in practice. Recycling plants, as shown in Table 1, were actually visited by the author from 2002 to 2007, and various assembly products were chosen. Based on interviews and the literatures, the plants are classified into the following: recovery circumstances, products, product and system characteristics (Table 1).

The recovery circumstances first show the recovery type (reuse or recycling), and products and the companies to which they belong. Then the plants are described. Depending on the products, they are regulated by Japanese recycling laws or Industry guidelines, and are identified in Table 1. EOL (End of Lifecycle) products have different processes based on product types, and are distinguishable by the postconsumer phase [4]. The details of the recovery circumstances are as follows:

Recovery type Based on [4], there are 2 types of product recovery: Reuse and Recycling. The definitions are described in an explanation of Postconsumer Phase.

Products In this survey, assembly products, which are used at home and at the office for the general consumer, are selected such as electronic products and automobiles. The products in surveyed factories are as follows: Recycling cameras, copiers, computers, TVs, refrigerators, washing machines, air conditioners, office equipment (copiers, faxes, etc.), cellular phones and automobiles.

Companies and plants Operating companies of inverse manufacturing and disassembly factories are classified by capital: 1) Dependent (Group), 2) Independent, 3) Joint capital.

Regulations Japanese recycling laws are already effective for electric appliances and automobiles. Consumers must pay the recycling fee when they buy the products, and when the EOL products are delivered to appropriate sites. On the other hand, industries initiative and operate their own common guideline to collect and recovery EOL products of computers and cellular phones.

Postconsumer Phase Based on [4], treatments of the EOL products in Postconsumer Phase are defined as follows: Refurbishing, Remanufacturing, Reuse, Recycling, Recovery, Cascading and Downcycling.

Product Characteristics

Inverse manufacturing and disassembly systems strongly depend on characteristics of products and systems. This is because the disassembly of EOL products is strongly constrained more than it is during their assembly. The product characteristics are described as follows:

Length of life cycle The life cycle here means a physical life cycle [4], which refers to the duration of time that spans from the production of a product up to the moment that it is discarded. During that time, a user purchases and uses a product, and finishes the usage. The length of the life cycle has a close relation with the obsolescence of product values and lifetime.

Obsolescence Products often have obsolescence of values [5] as time passes before the end of their physical lifetime. It is difficult to reuse products with the obsolescence of values because consumers do not buy them on the market.

Size and weight Product size and weight for each product can constrain inverse manufacturing and disassembly systems in view of configuration and material handling of the system.

Others Certain products should be carefully treated in inverse manufacturing and disassembly systems in operation when they have hazardous materials and/or personal information. For example, air conditioners have hazardous CFC materials which are necessary for appropriate and special processes, and computers and cellular phones store personal information, and in all cases, the data must be erased before discarding.

System Characteristics

Systems characteristics are different from product to product, and are basically restricted by size and weight in the product characteristics. The details are as follows:

Configuration System configurations are typically classified into Line, Cell and Flexible types as well as into assembly systems described in [7]. Line type traditionally consists of serial disassembly

stations connected by material handlings such as conveyors. The division of labor is carried out at each station, and balance losses occur. Input units at the system are received differently with services at all stations from the first to the last station. Cell one has the self-completion stations, and it does not have the balance losses, but is isolated among the stations and does not have the pace of work by the material handlings such as conveyors. This is because its stations are not connected by material handlings. Then, the Flexible Cell System (FCS), where the self-completion stations in an ordered-entry array are connected by conveyors, enables both efficient flow and processing of units.

Material handling The flow of the units is performed by material handling equipment, such as conveyors and AGV (Automated Guided Vehicle) or operators (manual). The handlings are often restricted by the size and weight of the products.

Disassembly process The processes, which are carried out in the plants, are described in this area. The typical processes are as follows: (Automated/Manual) Sorting, (Automated/Manual) Disassembly, Cleaning, Data Erasing, and Assembly.

ISSUES FOR REUSE

Recycling Cameras

Reuse is essential to construct inverse manufacturing systems, and its application is assumed in a significant amount of researches for closed-loop supply chains. With Reuse, there are 2 types [3]: product installation reuse and spare parts reuse. However, it has been found that there are still a few cases of product installation reuse in Japanese practice. Recycling cameras by FUJI FILM are well known, and their production and recovery employs an ideal inverse manufacturing system.

Selected parts are reused for product installation of new products in assembly systems. In view of product characteristics, the length of the life cycle is typically very short, less than 1 year, because the finished cameras are returned to a disassembly system from users in order to develop images at photo stores. Also, matured technologies are used to design and produce the products. Therefore, the obsolescence of the product/parts values is very low, though over 110 models have been developed since the first model was released. This enables reuse for selected and expensive parts such as flash, body unit, lens and battery as new products.

The other parts are processed for material recycling. In view of system characteristics, factory automations are built for almost all stations, especially for sorting and disassembly to save labor costs. During the sorting process, reverse blocking occurs [8][9]. However, the demand for recycling cameras is decreasing in Japan since the market for these kinds of cameras is drastically shifting from film to digital. Then, the break even point for profit is tight, and it is difficult to implement new equipment investment and depreciation of automated equipment. When demand decreases, manual operation can decrease the number of operators. However, automated operation is not an easy way of saving operating costs when demand decreases. One of the issues to be addressed is how to decrease production costs when demand and sales reward decrease in view of product lifecycle.

Copiers

Copiers by RICOH are also reused as recycled copiers, and RICOH UNITECHNO has an inverse manufacturing system by refurbishing/remanufacturing, reuse and recycling. The main models, which have a larger production quantity of about 5% in all models, are selected, refurbished, remanufactured, reused and recycled using reused parts and components as reconditioning machines.

Table 1 Japanese Cases for Inverse Manufacturing and Disassembly Systems

Recovery Circumstances				Product Characteristics					System Characteristics				
Recovery type	Products	Copmanies and plants	Regulations	Postconsumer phase	Length of lifecycle	Obsolescence	Size and weight	Others	Configuration	Matarial handling	Disassembly treatments		
Reuse	Recycling Cameras	Dependent		Reuse and Recycling	6 weeks to 1 year	little	small light		Automated Cell	Conveyor and AGV	Automated Disassembly Automated Sorting Production installation reuse in new products		
	Copy Machines	Dependent		Refurbish, Reuse and Recycling	2.5 to 10 years	some	Large Heavy		Free Flow Line	Manual	Diagnosis, Disassembly, Cleaning, Replace, Inspection		
	Computers	Dependent	Gudeline	Refurbishing	3 to 5 years by lease	vaule	medium	carefully users personal data	Cell	Manual	Delete Date, Cleaning, Disassembly, Assembly		
Recycling	TVs Refrigators	Dependent&Joint	Law	Recycling	10 years	large	large heavy	carefully CFC	Line Line	Conveyor Conveyor	manual disassembly manual disassembly		
	Wathing machines							for Air	Line	Conveyor	manual disassembly		
	Air conditioners							conditoner	Line	Conveyor	manual disassembly		
	TVs Refrigators	Dependent	Law	Recycling	10 years	large	large heavy	carefully CFC	Line Line	Conveyor Conveyor	manual disassembly manual disassembly		
	Wathing machines							for Air	Line	Conveyor	manual disassembly		
	Air conditioners							conditoner	Line	Conveyor	manual disassembly		
	TVs Refrigators Wathing machines Air conditioners	Independent	Law	Recycling	10 years	large	large heavy	carefully CFC	FCS FCS	Conveyor Conveyor Conveyor	manual disassembly manual disassembly manual disassembly		
	Office Automation Equipment			Independent		Recycling	10 years	some	Large Heavy		Cell	Manual	manual disassembly
	Cellular Phone			Independent	Gudeline	Recycling	6 months	large	very small light	carefully users personal data	Job Shop	Manual	manual disassembly
	Automobile	Independent	Law	Recycling and Spare Parts Reuse	10 years	large	very big very heavy		Line	Conveyor	manual disassembly		

The individual products are managed by information systems. Based on the limited number of copies, etc., the collected products are diagnosed for refurbishing/remanufacturing as reconditioning machines, parts/components are reused for the reconditioned machines, or for recycling.

Reconditioning machines from production installation are reused more than 80% based on weight, and are more effective than new machines installed with reused parts by several dozen % on LCA. This indicates that the difficulties to treat both reconditioned machines and new machines for reused parts in the same discussion.

Another issue to be addressed is difficulty for copiers to have several repeated life cycles by reuse. This is because the length of life cycle for one product/lifetime is very long, up to 10 years. Therefore, selected products only can have a 20 year life cycle as a new product and reconditioned machine. It may be difficult for a product 20 years or older to have the 3rd life cycle by reuse.

Computers

Computers have an advanced modularity design by standard connection among parts such as hard disks, monitors, CPU, memory, and they are refurbished as refurbished computers or parts. However, the obsolescence of the value for performance occurs so rapidly, and the repeated life cycles are difficult to design.

ISSUES FOR RECYCLING

TVs, Refrigerators, Washing Machines and Air Conditioners

TVs, refrigerators, washing machines, and air conditioners are restricted by home appliance recycling laws in Japan. When a product is disposed, consumers have to pay a recycling fee, and these EOL product are collected. In view of product characteristics, the length of the life cycle is 10 years, the obsolescence of the value occurs and the lifetime is finished. Therefore, reuse of parts is impossible, and material recycling is carried out. Those products have many metal materials, and the metal materials are derived from shredding after manual disassembly. Since the size and weight of the products are large and heavy, conveyors are introduced for material handling in disassembly plants. With TVs, a flexible cell system [7] is used. Recently, reduction of recycling fees for users has been implemented because market prices for recycled materials, especially metals, are increasing. Since manual disassembly is expensive, and the disassembly treatments are decided based on the prices for recycled materials. One issue is that the operation strongly depends on the market prices of recycled materials.

Cell Phones

A case of the cellular phone is introduced in a closed-loop supply chain [1]. In Japan the cellular phone has many functions such as uses as a telephone, a GPS, a digital camera and the internet (e-mail and web browsing). New models are basically released twice a year, and the length of the life cycle is short. However, the rate of collection for used products is low because each cellular phone has personal data including photos, e-mails and phone books, and the size of the phone is very small. Cellular phones are made of a lot of rare and expensive precious metals, and manual disassembly is carried out in disassembly plants. Recycling plants do not require special material handling since the products are very small and light. On the other hand, there is a possibility to design repeated life cycles for reuse, because the length of life cycles is very short. One issue to be addressed is how to design modularity for reusable parts.

Automobiles

Automobiles are basically disassembled for material recycling. Some parts in selected automobiles are reused for spare parts, but the rate is only 10%.

SUMMARY

This study focuses on inverse manufacturing and disassembly systems in actual Japanese cases, and discusses their system design by considering repeated product life cycles among product generations and types. The obsolescence of the values is essential for reuse, and the repeated life cycles should be designed for inverse manufacturing and disassembly systems. Further studies should be conducted to model the inverse manufacturing systems in view of repeated life cycles of products, etc.

ACKNOWLEDGMENTS

The factories and plants visited by the author are gratefully thanked for helping with this survey. This research was partially supported by the Ministry of Education, Science, Sports and Culture, a Grant-in-Aid for Young Scientists (B), 17710127 in 2007.

REFERENCES

- [1] Guide, V., Danile R. *Business Aspects of Closed-Loop Supply Chains*. Pittsburgh, PA, USA: Carnegie Mellon Univ. Press, 2003.
- [2] Gungor, A., Gupta, S. M. "Issues in Environmentally Conscious Manufacturing and Product Recovery: A Survey." *Computers and Industrial Engineering*, 1999, 36(4), 811-853.
- [3] Kobayashi, H. *Product Lifecycle Planning – a practice for ISO/TR14062 -*. Tokyo, Japan: Ohmsha, 2003 (in Japanese).
- [4] Lambert, A. J. D., Gupta, S. M. *Disassembly Modeling for Assembly, Maintenance, Reuse, and Recycling*. Boca Raton, Florida, USA: CRC Press, 2005.
- [5] Lifecycle Design Committee, Inverse Manufacturing Forum. *Inverse Manufacturing and Lifecycle Design –for a creation of a recycling society-*. Tokyo, Japan: Manufacturing Science and Technology Center, Maruzen, 2002 (in Japanese).
- [6] Seliger, G., Kernbaum, S., Zettle, M. "Moving Towards Sustainability in Engineering." *IV Global Conference on Sustainable Manufacturing, Development and Life Cycle Engineering*, Sao Carlos, Brazil, Oct 2006, on CD-ROM.
- [7] Yamada, T., Kakefuda, S., Matsui, M. "Strategic Selection of Assembly Systems under Viable Demands." *Assembly Automation*, 2006, 26(4), 335-342.
- [8] Yamada, T., Matsui, M. "Disassembly Production Systems and Its Design Issues." *Reprints of Japan Industrial Management Association*, Spring meeting, Chofu, Japan, May 2003, 144-145 (in Japanese).
- [9] Yamada, T., Yamamoto, H., Matsui, M. "A Queuing Model of Disassembly Systems with New Blocking." *Proceedings of PCMM' 2004*, Gold Coast, Australia, Dec 2004, 2, 778-785.

AN INVENTORY CONTROL MODEL FOR PRODUCT RETURNS

Badr O. Johar, Northeastern University, Boston, MA 02115, (617)-373-7635, johar.b@neu.edu
Surendra M. Gupta*, Northeastern University, Boston, MA 02115, (617)-373-4846, gupta@neu.edu
(*corresponding author)

ABSTRACT

In recent years, the sales of consumer electronics have increased rapidly. Manufacturers introduce the state of the art technologies in quick succession. As a result, the volume of the returned end-of-life (EOL) products has grown significantly [1]. Often these EOL products are still in excellent working conditions (functional). Customers return them because various marketing programs create an aura of “must have” sense in the minds of customers or because of the favorable incentives offered by service providers or manufacturers to acquire upgraded products. Thus, disassembly facilities are often faced with an overwhelming number of returns that is beyond the actual demand of these items. The appropriate management of the reverse supply chain, specifically in terms of inventory control, then becomes more complicated. In this paper, we address the issue of managing the inventory generated in the disassembly facility that is beyond the forecasted demand and propose an inventory control model to minimize its impact. An example is considered to illustrate the approach.

INTRODUCTION

The environmentalists have always demanded that companies should take end-of-life (EOL) products back and manage them in an environmentally conscious manner. EOL products can be remanufactured, reused, recycled, or disposed of. However, because of the uncertainty in quantity and quality of items taken back, planners have often felt discouraged to engage in such initiatives. Other factors that inhibit them from engaging in such activities include: less vertical integration, lack of efficient recovery process, and questionable profitability.

Due to much rigid environmental legislation and the understanding of economical and environmental benefits from minimizing the use of virgin resources, many corporations 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 such as limitations on waste disposal and recycling requirements. The idea behind it is to allow the use of materials and parts more than once before it is finally discarded [2]. Thus, effective supply chain management is vital in gaining a competitive edge over other corporations [3]. Supply chain management “involves functions such as production, purchasing, materials management, warehousing, inventory control, distribution, shipping, and transport logistics” [3]. However, the take-back process is much more environmentally friendly than the traditional supply chain process as it “closes the loop” of the supply chain process and transform the end of life (EOL) items into new serviceable products. This new portion of the supply chain is known as Reverse Supply Chain and together they are known as Closed-Loop supply Chain.

The management of reverse supply chain is challenging compared to regular supply chain because it is much more reactive and much less visible [4]. Examples include, changing demand rates, multiple demand sources, variety of products, and logistical complications. Yet another challenge is offered by inventory control and value management of EOL products. Due to the

disparity between demands for parts and materials and their yields, the planner faces economical constraints (holding and transaction cost of excess inventory) and physical constraints (capacity and space limitations) when trying to make an informative decision of how many products to take back and when to take them back. Our research focuses on developing an inventory control policy of an on-hand inventory of returned items and disassembled parts such that profit is maximized.

The returned EOL products are disassembled into separate parts and materials. For large quantities of returned products, the best way to disassemble them is to do it on a disassembly line that consists of several disassembly workstations [5]. Workstations tend to experience different accumulation rates as well as different depletion rates because of differences in yields and demands. Such differences create “uncertainties” in inventories and space requirements at the workstations. It is therefore necessary to develop a method to determine an appropriate inventory policy. In most cases, the disassembly system ends up with on-hand inventory in excess of what is needed. In this paper we obtain the appropriate pricing of disassembled parts and associated costs such as transaction costs, holding costs, and recovery costs. We develop a static approach where the liquidation time is predetermined and a dynamic approach that depends more on the number of items (quantity) on hand. We propose an inventory control policy that tracks the total accumulated cost of on-hand items and measure its effect on the system’s profit.

INVENTORY CONTROL IN REVERSE LOGISTICS

One of the most challenging issues in reverse logistics in general and in disassembly line in particular, is the inventory management of returned items and disassembled parts. In previous research we proposed a linear programming model to balance the inventory of returned items while satisfying the demand. Yet, the system still ends up with excess on hand inventory (OHI) of disassembled parts that is in excess of what is needed. The challenge is to determine the optimal liquidation batch sizes for each item type in conjunction with suggested OHI. In addition, the costs associated with the processes are different and should be compared and prioritized based on some appropriate criteria. For example, it might be more desirable to dispose of hazardous materials quickly (incurring some cost) rather than carrying them in inventory for a potentially large profit in the future but at a risk of incurring a greater loss due to the possibility of contamination of other parts [6]. Timings of such decisions are critical because all items kept in inventory incur holding costs and carry a risk of obsolescence. Other items could cost more to hold rather than their actual remaining values.

INVENTORY AND VALUE MANAGEMENT OF EOL

Problem Description

Products disassembly and recovery is a complex problem. One of the issues in the reverse supply chain is inventory control and management of EOL products. Returned products arrive to disassembly facilities at different rates. Demands for disassembled parts and subassemblies occur at different rates and at all workstations. Disassembly facilities have no control over the quantity, quality, and flow of these EOL products. These products have different quality levels depending on the different environments they have been operating in. The main goal is to satisfy the demand for these products and items from the EOL products returns. Any excess demand is outsourced from an outside supplier. Because of the disparity between the demand and the line yields this will create uncertainty of the inventory levels of items that should be kept on hand to

cover for any future uncertainty. Such disparity will create a “fluctuation” of the amount of inventory to keep on hand, and when to liquidate it. In previous research, we introduced a linear programming model that minimizes the total cost of the system while satisfying the demand. However, the system tends to accumulate the EOL products which are disposed of when they reach a critical level. Nevertheless, by discounting the prices on current inventory, one might be able to slow down the rate of accumulation. Yet, we are also faced with a demand source capacity regarding the quantity the customers can purchase and a storage space limitation at the disassembly facility. We illustrate the problem graphically in figure 1.

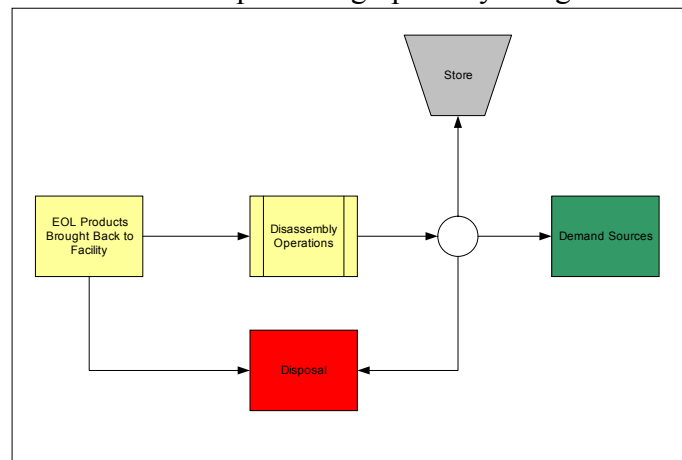


FIGURE 1

Graphical representation of the process

Problem Definition

After performing disassembly operations on returned EOL products, inventory of different items will start to pile up. Disassembled parts are being used to satisfy the demand and any shortage is being outsourced from an outside suppliers. In some case we will have an excess inventory beyond the demand. Appropriate management of this excess inventory is crucial to minimize the overall cost of the system and maximize its profit.

Problem Assumptions

1. There is a single product with three quality levels, and static deterministic inventory
2. Demand and returns are constant and known.
3. OHI is given and is different from period to period (Remaining after demand is satisfied)
4. There are zero inventories at the end of last period (end of planning horizon).
5. The price change is non-linear.

Problem Objective Function

The objective is to maximize the profit of the sub-system, which is revenue (money collected from the sales of the disassembled parts/subassemblies), minus the associated cost (transaction cost, holding cost, transferring-to-storage cost). The goal in this type of situation is to determine the minimum inventory we can hold before we decide to liquidate/dispose excess inventory.

METHODOLOGY AND PROBLEM FORMULATION

Notations

Known Parameters

- $C_{i,k,t}^{dc}$ Cost to dispose one unit of core type i , with quality level k , in period t
 $C_{j,t}^{diss}$ Cost to dispose one unit of subassembly j , with quality level k , in period t
 $C_{i,k,t}^{holc}$ Cost to hold one unit of core type i , with quality level k , in period t
 $C_{j,k,t}^{hols}$ Cost to hold one unit of subassembly type j , with quality level k , in period t
 $C_{i,k,t}^{trac}$ Cost of transaction associated with selling one unit of core type i , with quality level k , in period t
 $C_{j,k,t}^{tras}$ Cost of transaction associated with selling one unit of subassembly/ part type j , with quality level k , in period t
 $P_{i,k,t}^c$ Price of core product type i , with quality level k , in period t
 $P_{j,k,t}^s$ Price of part/subassembly type j , with quality level k , in period t
 $Q_{i,k,t}^{ohic}$ Quantity of on hand inventory of core type i , with quality level k , available in period t
 $Q_{j,k,t}^{ohis}$ Quantity of on hand inventory of subassembly/part type j , with quality level k , available in period t
 w_i Amount of space consumed by one unit of core type i
 w_j Amount of space consumed by one unit of subassembly/part type j
 CAP_n^{dem} Demand source n capacity in terms of cubic units
 CAP_i^{stgc} Storage facility capacity of core products type i
 CAP_j^{stgs} Storage facility capacity of subassembly/part type j

Decision Variables

- $Q_{i,k,t}^{liqc}$ Quantity of core product type i , with quality level k , liquidated in period t
 $Q_{j,k,t}^{liqs}$ Quantity of subassembly/part type j , with quality level k , liquidated in period t
 $Q_{i,k,t}^{remc}$ Quantity of core product type i , with quality level k , remaining at end of period t
 $Q_{j,k,t}^{rems}$ Quantity of subassembly/part type j , with quality level k , remaining at end of period t

Problem Formulation

$$\begin{aligned}
 \text{Max } R = & \sum_i \sum_k \sum_t P_{i,k,t}^c * Q_{i,k,t}^{liqc} + \sum_j \sum_k \sum_t P_{j,k,t}^s * Q_{j,k,t}^{liqs} - \sum_i \sum_k \sum_t C_{i,k,t}^{trac} * Q_{i,k,t}^{liqc} \\
 & - \sum_j \sum_k \sum_t C_{j,k,t}^{tras} * Q_{j,k,t}^{liqs} - \sum_i \sum_k \sum_t C_{i,k,t}^{holc} * Q_{i,k,t}^{remc} - \sum_j \sum_k \sum_t C_{j,k,t}^{hols} * Q_{j,k,t}^{rems}
 \end{aligned}$$

Subject to

$$Q_{i,k,t}^{liqc} + Q_{i,k,t}^{remc} \leq Q_{i,k,t}^{ohic}, \forall i, k, t \quad (1)$$

$$Q_{j,k,t}^{liqs} + Q_{j,k,t}^{rems} \leq Q_{j,k,t}^{ohis}, \forall j, k, t \quad (2)$$

$$\sum_i \sum_k w_i * Q_{i,1,t}^{liqc} + \sum_j \sum_k w_j * Q_{j,1,t}^{liqs} \leq CAP_1^{dem}, \forall t \quad (3)$$

$$\sum_i \sum_k w_i * Q_{i,k,t}^{liqc} + \sum_j \sum_k w_j * Q_{j,k,t}^{liqs} \leq CAP_2^{dem}, \forall t, k \neq 1 \quad (4)$$

$$\sum_i \sum_k Q_{i,k,t}^{remc} \leq CAP_i^{stgc}, \forall t \quad (5)$$

$$\sum_j \sum_k Q_{j,k,t}^{rems} \leq CAP_j^{stgs}, \forall t \quad (6)$$

All variables $\geq 0, \forall i, j, k, t$

NUMERICAL EXAMPLE

The inventory model in this paper is an extension model of our previous research [6]. Previously, we developed a linear programming model to simulate the accumulation of inventory on hand. Demand sources were: direct reuse, recycling, and disposal. We took in consideration the environmental regulations. In this model we introduced two demand sources: remanufacturing, which only takes good quality items, and refurbishing which takes average and/or imperfect quality items. The data and the results are presented in tables 1 and respectively.

Table 1
Data used for on hand inventory of EOL

Part/Subassembly Type	Quality Level	Space Req.	Profit/Unit	Holding C.	Transaction C.
Core Type 1	Good	0.625	\$37.00	\$3.70	\$1.50
	Average	0.625	\$23.00	\$2.30	\$1.50
	Imperfect	0.625	\$17.00	\$1.70	\$1.50
Part Type 1	Good	0.1	\$13.00	\$1.00	\$0.75
	Average	0.1	\$10.00	\$1.30	\$0.75
	Imperfect	0.1	\$8.00	\$0.70	\$0.75
Part Type 2	Good	0.2	\$8.00	\$0.50	\$0.75
	Average	0.2	\$4.00	\$0.55	\$0.75
	Imperfect	0.2	\$3.00	\$0.20	\$0.75
Part Type 3	Good	0.25	\$7.00	\$0.33	\$0.75
	Average	0.25	\$5.00	\$0.40	\$0.75
	Imperfect	0.25	\$2.00	\$0.13	\$0.75

The model was run over five periods and analyzed using Lingo 8 and a spreadsheet. Due to limited space, only results from period 4 are presented in table 2.

The objective function value is calculated to be **\$178,770.80**. The objective function value did not account for a small fraction of end of period inventory that could be either disposed of or transferred to a remote storage facility. Note that the system has a tendency of liquidating higher value items more often and keep lower valued ones on hand. This is because that demand can be satisfied by any quality level available, and by liquidating a higher value items faster it results in a reduction in carrying cost and an increase in the performance level of the system, and eventually the system's overall profit.

Table 2
Results from period 4

Period	Type	Quality	OHI	Liq	Rem	Acc. Inv	Dispose	Actual Inv.	Extra Inv.
Period 4	Core Type 1	Good	526	526	0	0	0	0	0
		Average	714	707	7	7	0	7	0
		Imperfect	456	0	456	456	456	650	194
	Part Type 1	Good	225	225	0	0	0	0	0
		Average	456	456	0	0	0	0	0
		Imperfect	125	125	0	0	0	0	0
	Part Type 2	Good	222	222	0	0	0	0	0
		Average	369	0	369	369	0	369	0
		Imperfect	222	0	222	222	222	750	528
	Part Type 3	Good	121	121	0	0	0	0	0
		Average	231	0	231	231	0	231	0
		Imperfect	156	0	156	788	38	750	594

REFERENCES

- [1] Lu, Q., Christina, V., Stuart, J. A. and Rich, T., "A Practical Framework for the Reverse Supply Chain", *Proceedings of the 2000 IEEE International Symposium on Electronics and the Environment*, San Francisco, California, pp. 266-271, May 8-10, 2000.
- [2] Gunter, K. L., "Inventory and Value Management in Demanufacturing Facilities", *Ph.D. Dissertation*, Michigan Technological University, Houghton, Michigan, 2004.
- [3] 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.
- [4] Tibeem-Lembke, Ronald S., and Rogers, Dale S., "Differences between forward and reverse logistics in a retail environment", *Supply Chain Management: An International Journal*, Vol. 7, No. 5, 271-282, 2002.
- [5] 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.
- [6] Johar, B. O. and Gupta, S. M., "Balancing Inventory Generated from a Disassembly Line: Mathematical Approach", *Proceedings of the 2006 SPIE International Conference on Environmentally Conscious Manufacturing VI*, Vol. 6385, Boston, Massachusetts, pp.13-21, October 1-3, 2006.

OPTIMAL ORDERING POLICY IN A DISASSEMBLY-TO-ORDER SYSTEM WITH 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 a variety of returned products are disassembled in order to satisfy the demand for specified numbers of components. The main objective is to determine the optimal number of take-back end-of-life (EOL) products for the DTO system that maximizes the profit. A wide variety of products and subassemblies are considered for disassembly in order to meet the customer's demand for the different components. Several factors are considered before disassembling any product. Some of the factors include the condition of returned products, the different number of suppliers offering EOL products, and the quantity discount rate offered by each supplier. Product conditions affect the product's yield (e.g., older products tend to have a lower yield for usable components). A variety of products with different conditions and prices from a number of suppliers are considered. Finally, suppliers offer quantity discounts in order to increase their competitive edge. We solve this problem by using a model that maximizes the revenue and minimizes the total cost including the cost of acquiring the EOL products, the cost of disassembly, and the cost of disposing excess products and components. When solved, the model provides an optimal ordering policy for the DTO system. An example is considered to illustrate the use of the model.

INTRODUCTION

In recent years, consumers, manufacturing companies, and governments have become aware of various environmental issues and difficulties associated with the escalating influx of end-of-life (EOL) products. Overflowing waste sites, increasing pollution levels, diminishing natural resources, mounting governmental regulations and rising consumer desires have forced manufacturers to become more environmentally conscious [1]. However, being environmentally conscious is driven more by its business value rather than just regulations because manufacturers welcome high profits together with the potential of projecting a good image in the community [2].

Recent studies and data from the US Environmental Protection Agency (EPA) show an increase in the amount of waste generated by residents, businesses, and institutions. In 2005, the amount of Municipal Solid Waste (MSW) generated in the US was more than 245 million tons which translates into 4.5 pounds of waste per person per day. This is the result of consumers demanding newer products with higher technologies. Because manufacturing technologies are rapidly changing, and consumers tend to upgrade their products to newer models, premature disposition of old products occurs even though they are still in very good operating conditions. Consequently, this leads to shorter lives for the products and added harmful waste. Therefore, it is essential to properly manage this problem in order to minimize the negative impact on the environment [3].

Because raw materials are depleting fast, harmful waste is constantly increasing, and landfills are filling up rapidly; manufacturers have thought of different ways to deal with this growing problem. Nowadays, manufacturers have many choices to manage EOL products. For example EOL products can be remanufactured, reused, recycled or disposed of. Remanufacturing, reusing and recycling products can help decrease the rate of depletion of virgin resources. Disposal, on the other hand, should be used as a last resort as it is harmful to the environment, increases pollution, and reduces the number of landfills. Consequently, several environmental regulations have been established which aim to restrict the number of disposed products, components and materials. This helps in conserving the environment and reducing the amount of raw materials used [4].

Disassembly is often the first crucial step in remanufacturing, recycling, and disposing. Disassembly is defined as the process in which products are separated into their components and materials using non-destructive, semi-destructive, or destructive operations. The main objective of this process is to minimize manufacturers' dependency on natural resources which in return reduces their high rate of depletion. Disassembly also allows us to use a retuned product more than once before it is discarded. This reduces the wastes sent to landfills. Additionally, the cost of reusing the materials is much less than the cost of new materials.

In recent years, the field of disassembly has gained a lot of interest. Few researchers have focused on the area of disassembly process and how it affects the environment, while others have focused on financial aspects of the field. For example, Gungor and Gupta [1] provide a comprehensive survey of issues in environmentally conscious manufacturing, and product recovery. Srivastava [2] presented a comprehensive integrated view of the published literature on all aspects and facts on Green Supply Chain Management taking a reverse logistics angle so as to facilitate further study, practice and research. The recent book by Lambert and Gupta [5] gives a good understanding of the general area of disassembly. Gupta and Kongar [6] present a disassembly-to-order algorithm that incorporates the disassembly and recovery of used products to satisfy a certain demand for products while achieving various financial and "environmentally benign" goals.

DISASSEMBLY-TO-ORDER SYSTEM

In a disassembly-to-order (DTO) system, a wide variety of products and subassemblies are purchased. Inventories are only kept at the product and subassembly levels. These are disassembled into individual components as customers demand for specific components. Therefore, the manager has two critical decisions to make: how many EOL products and subassemblies to purchase before the final demand arrives, and once the demand arrives, how to disassemble the EOL products and subassemblies to maximize the profit. However, there are a lot of uncertainties in the disassembly process that complicate the process. Some of the problems include the conditions of the retuned EOL products, demand uncertainty, limited supply of products from each supplier, quantity discount rates offered by the suppliers, manufacturer inventory control policies, system lead times, and components due dates.

PROBLEM DEFINITION

In this paper, we consider the disassembly-to-order (DTO) system where a variety of returned products are disassembled in order to fulfill the demand for specified numbers of components. However, there are several factors that we need to consider before disassembling the products. First, EOL products in the disassembly process are received in a variety of conditions and this leads to a lot of uncertainties in the process. For example, the yield of components from each product is unknown and this complicates the problem. Older products tend to have lower yields than newer products. Therefore, we do not know in advance how many products we need to disassemble in order to fulfill the demand for components. Second, EOL products are supplied by a number of suppliers and each supplier can supply a wide range of products with different prices and in various conditions. Additionally, there are capacity constraints which limit the number and variety of products that each supplier can offer. Finally, suppliers offer quantity discounts to increase their competitive edge. To solve this problem, we develop a model that maximizes the revenue and minimizes the total cost including the cost of acquiring the EOL products, the cost of disassembly and the cost of disposing excess products and components. When solved, the model provides an optimal ordering policy for the DTO system.

Formulation

The following notations are used in this paper:

- i = Index donating the Product, $i = 1, 2, 3, \dots, x$
- j = Index donating the Supplier, $j = 1, 2, 3, \dots, y$
- k = Index donating the Component, $k = 1, 2, 3, \dots, z$
- π = Total profit of selling components k
- TP_j = Total purchase from supplier j
- x_{ij} = Number of products i purchased from supplier j
- p_{ij} = Price of buying product i from supplier j
- s_k = Price of selling component k
- DR_j = Discount rate of supplier j
- DC_k = Disposal cost of 1 unit of component k
- Num_{kij} = Number of components k in product i from supplier j
- DAC_{kij} = Cost of disassembling component k in product i from supplier j
- $TDAC_k$ = Total disassembly cost of component k
- Yld_{kij} = Yield of component k in product i from supplier j
- D_k = Demand of component k
- Cap_{ij} = Capacity of supplier j for product i
- Q_k = Total quantity of component k

Objective Function:

$$\text{Max } \pi = \sum_k (s_k * D_k) - \left[\sum_j (1 - DR_j) * TP_j + \sum_k DC_k * (Q_k - D_k) + \sum_k TDAC_k \right]$$

Subject to:

$$x_{ij} \leq Cap_{ij} \quad (1)$$

$$TP_j = \sum_j \sum_i (p_{ij} * x_{ij}) \quad (2)$$

$$Q_k = \sum_k \sum_j \sum_i (Yld_{kij} * x_{ij}) \quad (3)$$

$$\sum_k Q_k \geq D_k \quad (4)$$

$$TDAC_k = \sum_k \sum_j \sum_i (DAC_{kij} * Num_{kij} * x_{ij}) \quad (5)$$

NUMERICAL EXAMPLE

We consider a case example to illustrate the application of the disassembly-to-order model. The objective of our case example is to maximize profit. We have 4 suppliers, 6 products, and 4 components. Each supplier offers 6 different types of products, each being a combination of 2 different components. The products from each supplier are priced differently and the conditions of the products differ from one supplier to another. This difference in product conditions will result in different yields for different components. Each supplier offers a different discount schedule which is based on the total dollar amount of products purchased from that supplier. Tables 1 through 5 summarize the input data used in the model to calculate the optimal purchase quantity of each product from each supplier. This information includes the demand of each component and the discount schedule of each supplier.

Table 6 is a summary of information regarding the 6 products from each supplier including the yield of components, the purchase price, the capacity, and the optimal purchase quantity. The model gives an optimal solution with a total profit of \$1,968.57 and discount rate of 10% from supplier 1, discount rate of 7.5% from supplier 2, discount rate of 18% from supplier 3, discount rate of 20% from supplier 4.

Table 1: Component Demand

Component	Demand
A	364
B	516
C	843
D	468

Table 2: Supplier 1 Discount Rate

Discount Schedules Supplier 1	Discount Rate (DR1)
0 < 400	0.0%
400 < 800	3.0%
800 < 1200	5.0%
1200 < 1500	8.0%
1500 < 1800	10.0%
1800 < 2500	13.0%
>= 2500	20.0%

Table 3: Supplier 2 Discount Rate

Discount Schedules Supplier 2	Discount Rate (DR2)
0 < 500	0.0%
500 < 1000	2.5%
1000 < 1500	6.0%
1500 < 2000	7.5%
2000 < 2500	11.0%
2500 < 3000	15.0%
>= 3000	21.5%

Table 4: Supplier 3 Discount Rate

Discount Schedules Supplier 3	Discount Rate (DR3)
0 < 200	0.0%
200 < 400	2.5%
400 < 600	4.5%
600 < 800	7.0%
800 < 1000	9.5%
1000 < 1500	13.0%
>= 1500	18.0%

Table 5: Supplier 4 Discount Rate

Discount Schedules Supplier 4	Discount Rate (DR4)
0 < 200	0.0%
200 < 400	2.5%
400 < 600	6.0%
600 < 800	7.5%
800 < 1000	11.0%
1000 < 1500	15.0%
>= 1500	20.0%

Table 6: Model Results for Product Purchase Quantity from each Supplier

		Yield of Component				Purchasing Price	Capacity	Purchase Quantity
		A	B	C	D			
Supplier 1	Product 1	0.289	0.959	0.000	0.000	\$4.50	100	97
	Product 2	0.000	0.000	0.801	0.331	\$3.50	152	151
	Product 3	0.570	0.000	0.000	0.732	\$5.00	352	1
	Product 4	0.000	0.487	0.000	0.755	\$4.25	332	4
	Product 5	0.471	0.000	0.733	0.000	\$3.75	132	49
	Product 6	0.000	0.320	0.600	0.000	\$3.25	112	112
Supplier 2	Product 1	0.225	0.817	0.000	0.000	\$3.50	100	97
	Product 2	0.000	0.000	0.798	0.326	\$3.00	87	87
	Product 3	0.529	0.000	0.000	0.750	\$4.25	128	127
	Product 4	0.000	0.507	0.000	0.793	\$5.00	175	0
	Product 5	0.443	0.000	0.769	0.000	\$3.50	72	72
	Product 6	0.000	0.416	0.797	0.000	\$3.75	122	122
Supplier 3	Product 1	0.185	0.731	0.000	0.000	\$3.00	225	142
	Product 2	0.000	0.000	0.849	0.272	\$2.75	101	101
	Product 3	0.508	0.000	0.000	0.686	\$3.00	203	129
	Product 4	0.000	0.570	0.000	0.778	\$5.00	82	79
	Product 5	0.432	0.000	0.910	0.000	\$4.00	52	52
	Product 6	0.000	0.395	0.844	0.000	\$4.75	87	87
Supplier 4	Product 1	0.191	0.915	0.000	0.000	\$5.00	254	8
	Product 2	0.000	0.000	0.907	0.309	\$3.75	65	65
	Product 3	0.515	0.000	0.000	0.719	\$4.00	103	103
	Product 4	0.000	0.525	0.000	0.613	\$3.25	100	32
	Product 5	0.367	0.000	0.751	0.000	\$2.50	62	62
	Product 6	0.000	0.483	0.827	0.000	\$5.50	103	103

CONCLUSION

In this paper, we evaluated the disassembly-to-order (DTO) problem where a variety of returned products are to be disassembled in order to satisfy the demand for specified numbers of components. The main objective was to determine the optimal number of take-back end-of-life (EOL) products for the DTO system that maximizes the profit. Our model had to take into consideration several factors including the condition of returned products, the variety of products from different suppliers and quantity discounts offered by suppliers. As a result, the model was able to determine the best combination of the EOL products to be taken back from each supplier for disassembly in order to satisfy the demand of each component and maximize the profit.

REFERENCES

- [1] 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.
- [2] Srivastava, S. "Green supply-chain management: A state-of-the-art literature review," *International Journal of Management Reviews*, Vol. 9, No. 1, 53-80, 2007.
- [3] Imtavanich, P. and Gupta, S. M., "Solving a Disassembly-to-Order System by using Genetic Algorithm and Weighted Fuzzy Goal Programming," *Proceedings of the SPIE International Conference on Environmentally Conscious Manufacturing VI*, Vol. 6385, 2006.
- [4] Imtavanich, 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 VI*, Vol. 5583, 147-162, 2004.
- [5] 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.
- [6] Gupta, S. M. and Kongar, E., "A Disassembly-To-Order System", *Proceedings of the 2001 International POMS (Sao Paulo) Meeting*, 331-338, 2001.

OPTIMAL ORDERING POLICY FOR PRODUCT ACQUISITION IN A REMANUFACTURING SYSTEM

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

(*Corresponding author)

ABSTRACT

This paper deals with the product acquisition problem and considers returned product quality. 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 considered to show the implementation of the methodology.

INTRODUCTION

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 [4]. 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) [5].

This paper deals with the product acquisition problem and considers returned product quality. 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. 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 returned products.

The system is composed of the state that denotes the inventory levels of two quality types 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.

Gungor and Gupta [4] 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 restrict 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.

FORMULATION

We formulate a product acquisition 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 returned products.

Figure 1 shows the product acquisition 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 acquisition 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

The state of the system is denoted by

$$s(t) = (I_1(t), I_2(t))$$

Production Policy

$$P_1(t) = \min \{D(t), [I_1(t)]^+\}$$

$$P_2(t) = \max \{0, D(t) - [I_1(t)]^+, I_2(t)\}$$

where $[x]^+ = \max\{0, x\}$.

$$I_1(t+1) = I_1(t) + O_1(t-1) - P_1(t)$$

$$I_2(t+1) = I_2(t) + O_2(t-1) - P_2(t)$$

Action spaces are shown by

$$K_1(s(t)) = \{0, \dots, I_{\max 1} - I_1 - O_1(t-1)\}$$

and $K_2(s(t)) = \{0, \dots, I_{\max 2} - I_2 - O_2(t-1)\}$.

Each action means that

$$k_1 = O_1(t), k_2 = O_2(t).$$

Transition Probability is

$$P_{s(n),s(n+1)}(k_1, k_2) =$$

$$\begin{cases} \Pr\{D(t) = d\}, s(t+1) = \{I_1(t) + k_1 - P_1(t), I_2(t) + k_2 - P_2(t)\} \\ \text{Otherwise} & , 0. \end{cases}$$

The expected cost is given by

$$r_{s(t)}(k) = \sum_{n=1}^2 (k_n a_n + c_n P_n(t) + h_n I_n(t)) + c_b \max\{0, D(t) - (I_1(t) + I_2(t))\}$$

Optimal equation is shown by the following;

$$g + v_i = \min_{k_1 \in K_1(i), k_2 \in K_2(i)} \left\{ r_i(k_1, k_2) + \sum_{j \in S} p_{ij}(k_1, k_2) v_j \right\} \quad (i \in S)$$

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.

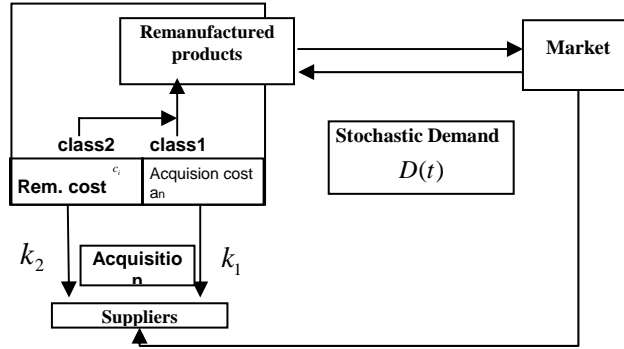


Fig.1: Product Acquisition Model

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

$$\Pr\{D_n = D - \frac{1}{2}Q + j\} = \binom{Q}{j} \left(\frac{1}{2}\right)^Q, (0 \leq j \leq Q)$$

where $D=3$ and Q is an even number and the variance(σ^2) is $Q/4$. We can obtain the expected average cost per period under the steady state of the system.

It is assumed that the maximum number of each inventory, $Imax_n=10$. The cost parameters are set as follows:

$$h_n=1(n=1,2), a_1=4, a_2=1, \text{ and } c_b=10.$$

We also assume that $C_1=1, C_5=1$ and $Q=3$. We can obtain the optimal policy using policy iteration method. Table 1 shows it under steady state with minimum cost, which is 28.94.

Table 1: Optimal ordering policy

$(I_1(t), I_2(t))$	k_1	k_2
(2, 2)	4	1
(3, 1)	3	2
(3, 2)	3	1
(4, 1)	2	2
(4, 2)	2	1
(4, 3)	2	0
(5, 3)	1	0

CONCLUSION

We formulated the acquisition problem together with product quality and stochastic demand using the Markov decision process in a remanufacturing system. Numerical results showed the optimal ordering policy that minimized the expected average cost per period for the product acquisition system with different kinds of quality classes.

ACKNOWLEDGMENT

This research is supported in part by JSPS Grants-in-Aid for Scientific Research No. 19510161.

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] Gungor, A. and Gupta, S. M., 1999, Issues in environmentally conscious manufacturing and product recovery: a survey. *Computers and Industrial Engineering*, 36, 811-853.
- [5] Guide, 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.
- [6] Gupta, S. M. and Taleb, K. N., 1994, Scheduling Disassembly. *International Journal of Production Research*, 32, 1857-1866.
- [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*. New York (John Wiley & Sons): Discrete Stochastic Dynamic Programming.
- [16] Wann, D., 1996, *Deep Design: Pathways to a Livable Future*. Washington: Island Press.

AIRLINE FLIGHT SCHEDULES, CYCLE TIMES, AND PROFITABILITY: 2004-2006

David West, East Carolina University, Greenville, NC 27858, 252-328-6370, westd@ecu.edu
John Bradley, East Carolina University, Greenville, NC 27858, bradleyj@ecu.edu

ABSTRACT

This research investigates the relationships between airline flight schedules, aircraft cycle times, and carrier profitability for ten large US domestic airlines. We find that point-to-point flight schedules and short cycle times are operational factors that airlines should exploit to improve profitability.

Keywords: cycle time, hub and spoke, operational strategy, profitability

INTRODUCTION

Two key operational factors that critically impact airline profitability are the design of a flight schedule defining a network of daily sources and destinations [Nero, 1999], and the operational capability to turnaround aircraft quickly [Gitell 2001, VanLandehan, 2002], (which we will refer to as the cycle time). Since deregulation of the US airline industry, the dominant flight schedule network of domestic carriers has been the hub-and-spoke (HS) strategy using multiple flight segments. With a HS network, a passenger is routed from an origin to a hub airport where the passenger connects to an outbound flight to the final destination. This strategy has been favored for its ability to aggregate passenger loads creating economies of traffic density, economies of scale and economies of scope. Toh and Higgins (1985) report weak evidence that a hub and spoke flight schedule increases airline profitability (based on operational data for 1982). Bania (1988) also reports empirical evidence that large multi-hub networks are an effective strategy for an airline to secure a competitive advantage.

Typical arguments for hub and spoke networks (i.e. Nero, 1999) cite efficiencies but overlook degradation of performance from congestion and capacity constraints at hubs and also overlook costs to the traveling passenger whose itineraries are significantly longer as they require two flight segments and extended layovers. Hub airports have become heavily congested with waves of arriving flights followed by waves of outbound flights scheduled at convenient travel times for passengers. As we intuitively expect (and this research confirms), the hubs require longer cycle times to turnaround aircraft. The extended cycle times undermine the human and physical productivity performance of the airlines and alienate customers. It is also noteworthy that airline revenue per passenger mile flown has been consistently declining over the past decade (Lee, 2003). At some point the airlines lose the ability to recover the costs of the two flight segments required by the HS network strategy. Revenue pressures are also aggravated by the increasing popularity of commuter jets which can efficiently serve point-to-point flight routing in smaller markets.

The purpose of this research is to investigate the relationships between airline flight schedules, aircraft cycle times, and airline profitability for large US domestic carriers. The results provide strong evidence that point-to-point-networks and short cycle times are vital to carrier profitability.

OPERATIONAL MODELS OF US AIRLINE INDUSTRY

Gittell (2001) argues that preparing flights for departure is one of the core processes of an airline's operations repeated hundreds of times daily in dozens of locations. The departure process is a set of interdependent tasks performed by a cross-functional group with 12 distinct functions: gate agents, ticketing agents, ramp agents, baggage handler, cabin cleaners, caterers, fuelers, freight agents, operations agents, pilots, flight attendants, and mechanics [Gittell, 2001]. The success or failure of the aircraft turnaround process can make or break an airline's reputation for both reliability and profitability by impacting productivity of the airline's most costly assets, employees and aircraft. Hult (2002) communicates the importance of speed as a competitive weapon and concludes that regardless of the nature of the industry, cycle time is a significant key to success. Despite the critical nature of cycle times in lean manufacturing systems, we find no published research linking cycle times and profitability in service industries including the airline industry. There are a number of research studies of the process of boarding passengers, a portion of the turnaround effort. For example, Van Landeghem (2002) recognizes that turnaround times for airplanes are under constant pressure to be reduced and provides a comprehensive simulation study of a number of potential boarding processes.

The design of a flight schedule for an airline is critically important; the flight schedule fixes a large portion of the airlines cost [Barnhardt, 2004; Dobson, 1993]. A common flight schedule strategy in the airline industry is the hub-and-spoke strategy (HS) which is defined by Rosenberger as a flight network with a large percentage of the flight segments into or out of a small subset of stations called hubs. A number of studies conclude HS networks allow airlines to improve economic returns, by exploiting economies of scale, economies of scope, flight frequency and enhancing market power at dominant hub locations [Brueckner and Zhang 2001, Adler, 2007, Nero, 1999, Barla & Constantatos, 2000, Dobson, 1993, Gillen, 2005]. The economic rationale for HS network is that the airlines can aggregate demand, increase frequency, reduce airfares and prevent entry of competitors into the marketplace (Adler, 2001). In a prior study of airline profitability conducted with data from 1982, Poh and Higgins report a strong and positive relationship between hub-and-spoke network centrality and airline profitability (but report an R2 of only 0.1).

There have been a number of studies that focus on the advantages of HS network flight schedules. For example, Adler 2001, uses a two-stage Nash best-response game, to evaluate the most profitable HS network for an airline to survive in a deregulated environment (Adler, 2001). Barla and Constantatos 2000, investigate the role of demand uncertainty in the adoption of the hub-and-spoke network structure in the airline industry. They show when an airline has to decide on its capacity before the demand conditions are perfectly known, a HS network structure helps the firm to lower its cost of excess capacity in the case of low demand and to improve its capacity allocation in the case of high demand by pooling passengers from several markets into the same plane (Barla & Constantatos, 2000). Button 2002, specifically highlights HS as protective of the airlines that center their operations on large hubs (Button, 2002). Nero, 1999, explores the extent to which airlines operating large hub-and-spoke networks secure a competitive advantage. His result shows that, "although the model exhibits decreasing returns to firm/network size, nonetheless there is a competitive advantage to increasing the size of a network" [Nero, 1999]. Marianov, Serra et al. 1999, investigate hub location on a network in a competitive environment. In their model, "each origin-destination flow can go through either one

or two hubs, and each demand point can be assigned to more than a hub, depending on the different destinations of its traffic” (Marianov, Serra, & ReVelle, 1999). Kelly and Bryan, 1998, purport an improved model for airline hub location by including the economies of scale resulting from the agglomeration of passenger flows. They show that current hub location models, by assuming flow-independent costs, not only miscalculate total network cost, but may also erroneously select optimal hub locations and allocations (Kelly & Bryan, 1998). Sasaki, Suzuki et al, 1999, consider the 1-stop multiple allocation p-hub median problem. They formulate the problem as a p-median problem and propose a branch-and-bound algorithm and a greedy-type heuristic algorithm and their results show that the proposed algorithms work better than the well-known nested-dual algorithm, particularly for relatively small problems (Sasaki, Suzuki, & Drezner, 1999).

Southwest is one airline that does not use a HS strategy in their flight schedule; instead they use a point-to-point (PP) network where a high percentage of passengers are routed directly from origin to destination. Using a game-theoretical framework for two carriers, Alderighi, Cento et al. 2005, show that asymmetric equilibrium may exist when carriers compete in designing their network configurations and that HS, PP and multi-hub strategies may coexist. They find “two main stable outcomes, which depend on the size of the internal market. First, when the internal markets are small, PP network strategies are played by both carriers, while for a specific subset of parameters a collusive equilibrium in a HS configuration can be derived. Second, when the size of the internal markets is large, asymmetric configurations, where one carrier chooses a HS strategy and the other chooses a PP strategy, are the only stable equilibrium” (Alderighi, Cento, Nijkamp, & Rietveld, 2005).

DATA DESCRIPTION AND EXPERIMENTAL METHODOLOGY

The data used in this study consists of 11.9 million domestic flights for ten major US carriers in 2004-2006; it is available from The Office of Airline Information of the Bureau of Transportation Statistics. This database contains a table of scheduled and actual departure and arrival times reported by certified U.S. air carriers that account for at least one percent of domestic scheduled passenger revenues (www.transtats.bts.gov). The database also contains domain tables that provide the acceptable values for many of the attributes. The primary table contains departure delays and arrival delays for non-stop domestic flights by major air carriers, and provides such additional items as origin and destination airports, flight numbers, scheduled and actual departure and arrival times, cancelled or diverted flights, taxi out and taxi in times, air time, and non-stop distance. The data is provided in monthly datasets that we downloaded and transformed into SQL Server tables. The primary table contained 20.6 million records for 2004-2006. Since Excel will only accept 65,000 rows, a program was written that opened the .csv-formatted table and processed it record at a time, transforming the data and storing it in an SQL table. Using SAS 9.0, this SQL table was defined as a data set for input into our analyses. From this data we calculated cycle times for each flight. We define cycle time as the time elapsed from wheels down when an aircraft touches down on arrival to wheels off when the aircraft departs.

The ten major US carriers investigated in this study are identified in Table 1. Commuter airlines, international flights, and charter operations were excluded from this analysis. The metric used to measure the relative reliance on HS flight design is the percentage of an airline’s routes that originate from the two most frequently used airports. This metric is also used by Bania (1998) to

measure the degree of centralization in route structure. Scale and economies of scale measure by number of airports served and number of direct flights.

RESULTS

The 13 most frequently used hub airports are identified in Table 1 based on their flight histories from 2004 to 2006. The average cycle time (wheels up minus wheels down) at the hub airports for the ten major U.S. domestic carriers is reported in Table 1. The hub with the longest cycle time is Houston where flights of domestic carriers averaged 130.5 minutes on the ground for each aircraft turnaround. Seven of the 13 hubs had cycle times in excess of 100 minutes including Minneapolis, Charlotte, Detroit, Atlanta, Chicago, and Dallas Fort Worth. The average cycle time for the 13 hub airports is 97.6 minutes; while all other major airports average 72.2 minutes. (In a later discussion we estimate the breakeven cycle time for an airline to be approximately 85 minutes). This difference of 25.4 minutes of cycle time between hubs and non-hubs is a statistically significant difference with $p < 0.001$. This cycle time penalty at hubs increases the cost of consolidating passengers. We value this penalty at \$457,000,000 per year for the average domestic carrier based on an opportunity cost estimated later in this section.

Table 1 Hub Strategy

Airport	Major Carrier	Number Departures/Year	Average Cycle Time (minutes)
Houston	Continental	231,311	130.5
Minneapolis	Northwest	299,773	117.9
Charlotte	US Air	249,677	115.6
Detroit	Northwest	297,446	113.5
Atlanta	Delta, Air Tran	663,958	106.1
Philadelphia	US Air	241,074	103.0
Chicago O'Hare	American, United	500,994	102.5
Dallas Fort Worth	American	465,584	100.7
Los Angeles		290,024	89.7
Denver	United	292,507	87.3
Phoenix	America West	407,643	86.0
Kansas City		239,866	73.6
Las Vegas	America West	338,264	68.9
Baltimore	Air Tran	224,688	60.1
Washington			
Total Hub		4,742,809	97.6
Total Non-Hub		10,374,013	72.2

The ten carriers investigated can be grouped into HS, PP and multi-hub network strategies. We characterize a single-hub HS network by a profile with a significant concentration of flights at the one hub and all other sources near or below 5% (i.e. Air Tran with 38% of flights at Atlanta). Other airlines that exhibit single-hub strategies include Alaska, ATA, Delta, and Jet Blue. The average cycle time for the 5 carriers employing the single hub HS network strategy is 88.1 minutes. An example of a 2-hub network strategy is American Airlines with hubs in Dallas Fort Worth, and Chicago accounting for 45% of American domestic flights. Other carriers with 2 hub networks include Continental, America West, North West, United, and US Air. The average cycle time for the 6 carriers employing the multi-hub HS network strategy is 100.1 minutes. By contrast, the two busiest airports for Southwest constitute only 7% and 6% of Southwest flights respectively. Southwest, with a cycle time of 46.5 minutes, is the only carrier of the twelve major airlines studied with a PP network strategy.

Our first empirical experiment is to relate the concentration of flights at hub airports to the airlines cycle time. The major hub airports have become heavily congested as air traffic has grown. The tendency to schedule flights in banks of arriving and departing aircraft during peak travel times also increases the congestion and delays, resulting in longer lead times. The consequence is that aircraft are on the ground longer times, which decreases aircraft utilization and the productivity of human and physical assets. The regression model of carrier cycle time on two-airport concentration (as reported in Table 1) is given in Figure 4. The PP network strategy of Southwest Airlines is clearly evident with a concentration of 0.13 and a cycle time of 46.5 minutes. The HS and multi-hub carriers form a cluster with higher concentrations and significantly longer cycle times. The regression model has a slope of 136.7 (p=0.0026) confirming that higher HS concentration is associated with longer cycle times. An increase in concentration of 0.10 achieved by scheduling more flights through dominant hub airports, increases aircraft cycle time by an average of 13.67 minutes per flight.

$$\text{Cycle time} = 139.2 * \text{Hub Concentration} + 30.2 \quad (R^2 = .708) \quad (1)$$

Table 2: Carrier Cycle Times and Profitability

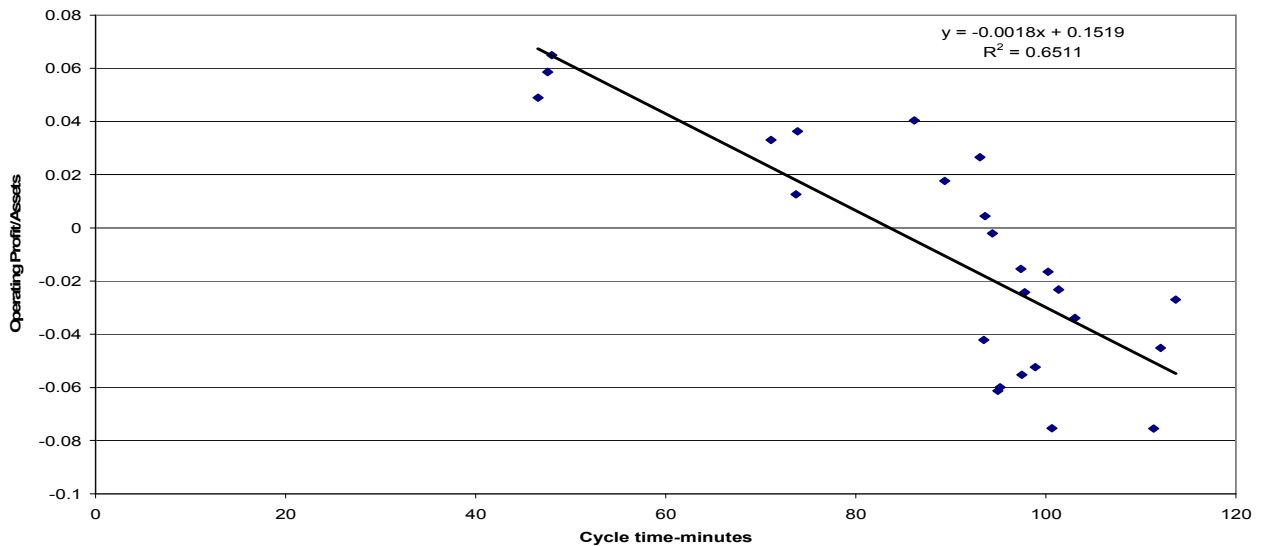
	Number	Percent	Cycle Time	Op- Profit/Assets
Southwest	794,507	23.0	46.54	0.049
Air Tran	124,122	3.6	73.86	0.036
Alaska	119,473	3.5	81.71	-0.012
Jet Blue	58,485	1.7	86.13	0.040
US Air	305,649	8.9	93.44	-0.042
America West	127,271	3.7	94.37	-0.002
United	362,845	10.5	97.45	-0.055
ATA	50,710	1.54	98.23	-1.750
American	444,233	12.9	100.23	-0.017
Delta	472,938	13.7	100.62	-0.075
Northwest	364,862	10.6	101.34	-0.023
Continental	183,834	5.3	113.64	-0.027

We isolated the financial performance for each carrier's domestic flight segments and calculated a ratio of the operating profit or loss divided by the assets employed in domestic operations. This

ratio is given in Table 2 where the carriers are listed in ascending order of cycle time. The regression of the ratio operating profit/assets on carrier cycle time is given in Figure 1. Note that ATA with a ratio of -1.75 is excluded as an outlier; we have not yet isolated the nature of the problem for this carrier. The estimate of the slope is significantly different from zero and estimated to be -0.0016 ($p=.012$). This confirms that increasing cycle time is associated with declining carrier profitability. A breakeven point for this model is a cycle time of 88.6 minutes. Carriers with faster cycle times are predicted to be profitable and conversely slower cycle times are unprofitable.

Figure 1

**Carrier Profitability vs Cycle Time
2004-2006**



CONCLUSIONS AND IMPLICATIONS

The results of this research are based on 11.9 million domestic flight records for calendar year 2004-2006. The data for this period reveals that the scheduling design decision to aggregate passengers through a hub network results in significantly longer aircraft cycle times. We also find that longer cycle times undermine the profitability of the carrier. The financial significance of cycle time reductions can be quantified from our regression models; the slope of -0.0018 represents the amount that the average carriers operating profits per dollar of assets invested will change with a one minute increase/decrease in cycle time. The average domestic airline has approximately \$10 billion in operating assets; a reduction of one minute in cycle time will, on average, increase operating profits by \$18 million.

The airlines could also benefit by reassessing strategic decisions in the schedule design of their network. Scheduling more direct flights and decreasing their reliance on congested hubs will increase operating profits. A reduction in hub concentration of 0.10 reduces average aircraft cycle time by 13.67 minutes which results in an estimated \$246 million increase in operating profits.

References are available upon request from David West

ENVIRONMENTAL REGULATION WITH GREEN TECHNOLOGY INNOVATION USING SUPPLY CHAIN INTEGRATION

Cheng-Wen Lee, Dept. of International Trade, Chung Yuan Christian University
Tel: +886-3-2655215, E-mail: chengwen@cycu.edu.tw

ABSTRACT

This paper examines the effect of environmental regulation on green technology innovation through the supply chain integration, using a multi-industry sample of manufacturing organizations. Building from the literature, this study establishes the research framework consisting of seven factors -- environmental regulation, market opportunity, external competition pressure, resource-based supply chain integration, network embeddedness, technology-related capacity, and green technology innovation. This study employs regression analysis and structural equation model to verify the eight hypotheses. Overall, the findings support all hypotheses but the last, H₇; that is, network embeddedness is not directly related to green technology innovation. Although environmental regulatory influence may vary across industries, this study suggests that environmental regulation does contribute to the enhancement of green technology innovation.

Keywords: Environmental regulation, Supply chain, Network embeddedness

INTRODUCTION

Because the adjustment of environmental regulations usually over time reflects updated understanding or new circumstances, there is a scope of strategic behavior. A firm may have incentive to strategically alter investment to induce favorable shifts in future environmental policy. A growing literature on dynamic environmental regulation includes related discussion of investment in new technology and learning (Jaffe, Newell and Stavins, 2003; Tarui and Polasky, 2005).

A number of studies regard environmental regulation from a strategic point of view, examining the relationship between environmental regulation and organizational performance (Shaffer, 1994; Dean and Brown, 1995; Sanchez and McKinley, 1998). An important aspect of organizational performance is "innovation", defined as the adoption of a device, system, policy, program, process, product or service that is new to the adopting organization (Damanpour, 1991; Collins, Hage and Hull, 1988).

Some of these studies imply a negative relationship, arguing that regulatory constraints add sizable compliance costs to firms, forcing cutbacks in R&D efforts and limiting innovation (Bartel and Thomas, 1987; Economist, 1996). Other studies argue that environmental regulation may appear disruptive and threatening to a firm, but it may actually be necessary to induce innovation (Schon, 1971). Whether environmental regulations inhibit or promote product/technology innovation seem at least partially dependant internal features of an organization (Sanchez and McKinley, 1998). Regardless of industrial views, economic activity produces a steady degradation of the environment that has increased pressure on firms to "green" their operations. The companies face mounting pressure to be accountable for their environmental management practices (Henriques and Sadorsky, 1999).

LITERATURE AND HYPOTHESE

Effect of environmental regulations

It is a common belief of the 1990s that differing environmental regulations can affect the competitiveness of industries and even countries. The extent to which environmental problems influence trade, and vice versa, has been a subject of considerable debate, although firms recognize the importance of efficient environmental regulations (Jayadevappa and Chhatre, 2000;

Beghin and Potier, 1997). The firms need to comply with environmental regulations while engaging in international trade and competing against those ones in the global market. In other words, the characteristics of markets in which firms sell products influence the decision-making of their environmental policies.

Business activities involve environment impact in the whole product life cycle, from raw material selection/exploitation, through production, packing, transport distribution, and installation maintenance to use and the end of life (disposal). Product life cycles have become shorter and large volumes of discarded products have contributed to the problem of electronics waste (e-waste). Such abandoned electronics, containing toxic substances, goes to landfills and contaminates the environment (Sarkis, 1995). Gangadharan (2006) found that firms that have a large percentage of sales to more developed countries like U.S., Canada and Europe have a higher probability of complying with environmental regulations. In particular, with the establishment of the WEEE and RoHS directives in the European Union, environmental standards for products are becoming more and more stringent. Companies are required to have a stronger sense of social responsibility with respect to environmental preservation. Consequently, Taiwan's firms also have to make green technology breakthrough and progress, as this is the gateway for international trade into Europe. Hence,

Hypothesis 1: Environmental regulation provides an incentive for accessing the green certificated market; that is "market opportunity". The greater the firms' compliance with environmental regulations, the greater their market opportunity will become.

Hypothesis 2: Environmental regulation increases external competition pressure among firms. Accordingly, the firms are compelled to improve green technology innovation for survival.

Green supply chain with resource-based perspective

Concerning the use of resource-based theory to predict technology innovation strategy, the focus is on which type of resources lead to an evolution in the product or process. That is, firms tend to develop new products and enter new markets where the resource requirements match their resource capabilities. Chief executive officers evaluate "firms' capabilities" as the resources that can represent sustainable competitive advantages (Snow and Hrebiniak, 1980). Chatterjee and Wernerfelt (1991) stated that intangible assets such as important strategic resources are relatively inflexible and motivated by marketing and innovative skills. Porter (1991) argued that aspects of dynamism relating to the growth of a firm may be overlooked as long as the firm has environmental opportunities. With respect to the outward focus, environment-related market opportunities are not explicitly incorporated in the empirical research pertaining to the boundaries of the firm (Andersen and Kheam, 1998). However, this study recognizes environment-related market opportunities as important factors in strategy analysis. If firms are unable to comply with environmental regulation, they may lose the opportunity to enter into the green certificate markets.

Green supply chain management (GSCM) has become a hot topic for electrical equipment and electronics manufacturers. Hu and Hsu (2005) examine critical factors for GSCM in Taiwan's electrical equipment and electronics industry. Many international companies such as Sony, HP, Dell and Apple have already established their green supply chains. Over the past decade there has been a worldwide trend for manufacturers to choose their green suppliers in the construction of green supply chain. Another hot topic is EU directives, including WEEE, RoHS, and EuP. These kinds of directives may influence several industries in Taiwan. Some studies focus on how the manufacturers comply with new directives and handle these affairs (Yang, 2005; Liu, 2006). The main objective of GSCM is to manage material cycles in the supply chain in an environmentally, socially, and economically responsible manner. Normally, this means that a product must generate as little waste as possible and conserve energy at each stage of the product's life cycle (Gramer,

1996). Consideration of this issue clearly demonstrates the growing importance played by the environmental dimension in the overall supply chain. It also highlights that companies need to modify their relationships with suppliers, shifting from competitive to co-operative behavior aimed at integrating the vendors in the green product/technology innovation process. Suppliers have to reduce the quantity of supplied components with low environmental performance and, hence, to comply with environmental regulations.

Green topics have an impact on the entire supply chain. Enterprises should not only look for alternative materials, but also place importance on the training and education of existing suppliers regarding environmental knowledge and practices. However, these environmentally friendly practices may induce the additional cost that it is possible to weaken the supply chain. Considering a solution for sharply reducing cost and slashing the wastes of resources, this study suggests that firms should pursue supply chain integration from a resource-based perspective.

Hypothesis 3: Desire increasing the market opportunity of entering into the green certificated market (e.g., EU), will simulate firms to adopt resource-based supply chain integration.

Hypothesis 4: Facing intense external competition upon entering the green certificated market (e.g., EU), firms will increasingly emphasize resource-based supply chain integration to enhance their competitive advantage.

Network embeddedness and technology-related capacity

Jarillo (1988) argued that networks link a chain of either inter-organizational technology or economic power to coincide with the strategic goal of improving intra-network participant's competitive advantage vis-a-vis market-place competitors. A kind of hierarchical relation exists between core companies and others of the network, and it involves transaction activities in the market for decreasing transaction cost such as reducing bargaining and monitoring costs (Porter and Fuller, 1987; Oliver, 1990; Kanai, 1992; Uzzi, 1996), but as a long-term relationship of unspecified contracts rather than as a structural task (Jarillo, 1988). Networks are vital to the discovery of opportunities, testing of ideas, and garnering of resources for the formation of the organization (Aldrich and Zimmer, 1986). In the context of organizations involved in economic transactions with several other organizations, Granovetter (1985) articulates the sociological concept of embeddedness. Potential partners are often very reluctant to put their reputation, capital, or other resources at risk in a start-up whose financial prospects, if not its longevity, are uncertain. "Embedded" constitutes ties with partners, "reinforced by mutual feelings of attachment, reciprocity, and trust" (Uzzi, 1996), that can enhance support for a start-up via the commitment of resources.

Each firm maintains unique and idiosyncratic patterns of network linkages and consequently is differentially exposed to new knowledge, ideas, and opportunities (McEvily and Zaheer, 1999). In reality, this differential exposure has been put forward as the basic competitive advantages of the firm, because it increases the breadth and variety of its network resources (Malnight, 1996). The notion of embeddedness originates from Karle Polanyi. According to Polanyi (1944), different economic forms differ from certain structural and institutional conditions. The dominant forms of integration in primitive and archaic societies (reciprocity and redistribution) are embedded in non-economic relations, like networks of culture or politics, religious, moral or judicial sanctions, and interpersonal relations (Polanyi, 1944). Embeddedness -- identified as the nature, depth, and extent of an individual's ties into the environment -- has recently been commented upon as a configuring element of general business process (Wittington, 1992; Uzzi, 1997; Dacin, Ventresca, and Beal, 1999). Halinen and Törnroos (1998) defined embeddedness in terms of "firms' relations with, and dependence on, various types of networks."

The application of the concept of network embeddedness to managerial issues is likely to become more prevalent (Oliver, 1990; Robinson, 2003; Jack and Anderson, 2002; Zeki, Perry, and Bartels,

2003). Dacin et al. (1999) recognize that network embeddedness can be looked upon as a strategic resource influencing the firm's future capability and expected performance. However, expected or perceived performance may vary between firms because of differences in network embeddedness. Most scholars propose that network embeddedness in business networks is assumed to develop over time from a state characterized by arm's length relationships to relationships based on adaptation and trust (Larson, 1992; Håkansson and Snehota, 1995; Ford, 1997). Consequently, network embeddedness should be treated as a continuous variable rather than as a dichotomy that is either absent or present (Dacin et al., 1999). Brain (1997) demonstrated that network embeddedness can provide economic benefits such as shortening lead-time, comprehensive understanding, and *pareto* resources distribution. In addition, the shared conceptual context provides motivation or incentive for sharing knowledge and information between the network members (Dyer and Nobeoka, 2000).

Technical embeddedness is one form of network embeddedness. Technical embeddedness is the interdependency between firms in terms of their product and production development processes. A high degree of technical embeddedness means that the two organizations are highly interdependent in terms of their technological activities (Andersson et al., 2002). There may be a closer relationship between alliance partners seller with involvement of R&D activity. This relationship may strength the ability to absorb new technological competence from the supply chain integration. It is reasonable to assume that activity links between firms related to technology are decisive factors behind the firm's capacity to adapt itself to new technology. Therefore,

Hypothesis 5: Through the resource-based supply chain integration process, it will positively influence the firms' network embeddedness.

Hypothesis 6: Using resource-based supply chain integration, firms will positively and significantly strengthen their technology-related capacity.

Green technology innovation

The development of environmental protection is strongly connected to the development of technology. Many critics argue that environmental regulations undermine innovation, and that complying with environmental regulation required by bureaucracy restricts firms from pursuing cutting-edge technology (Breyer, 1982). However, a dissenting argument is that environmental regulation, if viewed as an external jolt, can stimulate innovation within an organization (Marcus and Weber, 1989).

The creation of conditions for competition between exhibited technology and green technology reflect all private and social costs. Green technologies, which like any new technology have to compete with established technologies, remain in an unfavorable position. When the new products with green technologies arrive on the market, they have not reached their optimum performance in terms of cost and reliability. R&D programs and investment subsidies are the first methods introduced to improve performance and stimulate the diffusion of green technology (Menanteau, et al., 2003). For those technologies which are now nearly competitive, more specific instruments are used which aim at integrating renewable energies in the electricity generation market. The creation of green certificate markets provides an incentive to innovate green technologies. Producers of renewable electricity are under the constant pressure of competition because of the green certificate. This pressure creates incentives for potential investors not only to control the cost of equipment, but also to control operating costs once the equipment is installed (Menanteau et al., 2003). Competition among firms frequently involves product innovation and subsequent promotional activity. This study supposes that,

Hypothesis 7: Accumulated Network embeddedness drives green technology innovation.

Hypothesis 8: Strengthened technology-related capacity drives green technology innovation.

RESULT

Sample

Our review of the literature provides us with many indicators of corporate environmental management. In addition to designing a structured questionnaire, this study interviews senior managers -- plant managers, heads of production, and project managers -- of four categories of manufacturers. Data is collected from a random sample of 1,567 Taiwan's firms cited in the Commerce Industrial Services Portal. By February 2006, we complete 384 questionnaires. At that time a second questionnaire is sent out, leading to a total of 538 questionnaires. The rate of valid response is 34.3%. This study uses an ANOVA to test for sample bias between questionnaires returned in the first and second rounds, but finds no significant differences.

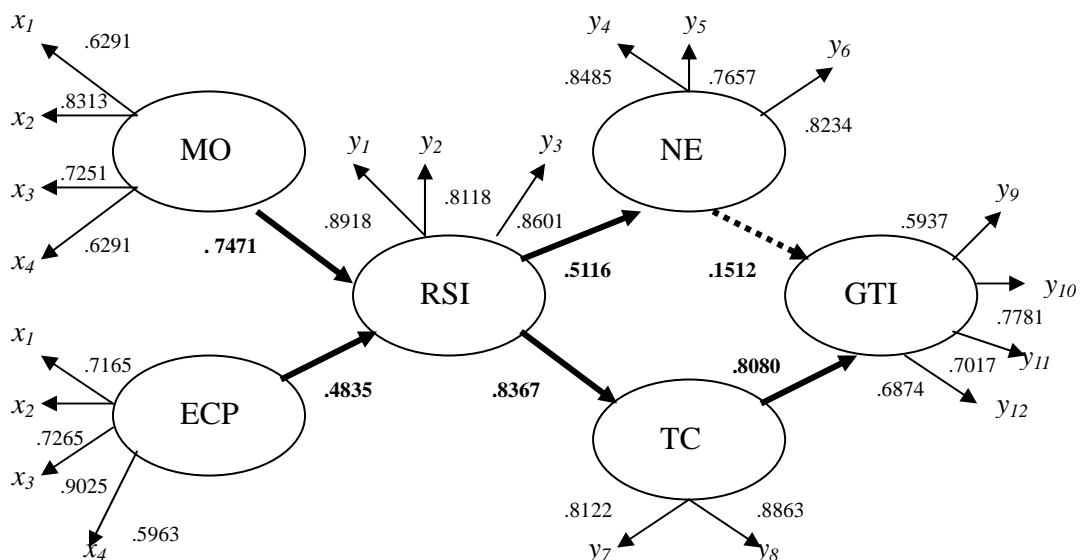
Structural model results

Table 1 presents the standardized SEM coefficients for each industry. Not all statistically significant coefficients in the four industry classes move in the expected direction. The study of measurement invariance in empirical research can also be initialized by simultaneously fitting to the covariance matrix the congeneric model in all groups under consideration. Imposing in it the cross-group equality of factor loadings and evaluating the resulting change in fit would similarly inform about group invariance in the underlying units of measurement (Byrne, 1998). Given the measurement invariance, this research combines the finding results of surveying four industries as Figure 1.

Table 1 SEM coefficients for each industry.

Equipments Path	Household equip.	Information/ communication equip.	Consumer electrical/ electronics equip.	Toy/ leisure/ sports equip.
MO→RSI	.690**	.735**	.801**	.700**
ECP→RSI	.538**	.455**	.538**	.454**
RSI→NE	.057	.347**	.612**	.375**
RSI→TC	.798**	.972**	.851**	.877**
NE→GTI	.142	.216**	.231*	.137
TC→GTI	.740**	.732**	.825**	.645**
Number of respondents	114	143	126	155

Level of significance: * p<0.05, ** p<0.01



Chi-square/df=1.863; GFI=.943; AGFI=.921
 CFI=.935; NFI=.902; SRMR=.018; RMSEA=.037

Figure 1 Structural model results (overall)

CONCLUSION

The findings of H₁ indicated that manufactures already profoundly comprehend deleterious substance stipulations; this is necessary for them to retain their market. Compliance with environmental regulations may allow them greater opportunity to access green certificate market such as the European Union. Regarding H₂, the other main driver of new practice adoption for compliance to environmental regulations is the so-called “bandwagon effect.” Bandwagons are diffusion processes whereby organizations adopt innovations because of external pressure caused by large numbers of other organizations that have already adopted the concept (Tolbert and Zucker, 1983). This survey support H₃. Supply chains not only allow firms to better satisfy existing clients, but also to win over the most profitable customers in new markets. As observed by the support for H₄, competition among firms frequently involves strategic integration in order to acquire counter-partners’ resources.

The findings of H₅ found that the stronger the network embeddedness, the greater the technology-related capacity created between network members. The result of H₆ confirmed Ahuja and Katila’s (2001) findings that integration significantly influences firms’ technological capacity and innovation performance while involving technological components. Moreover, according to firm resource-based views, firm-specific assets housed within one organization are merged with assets in another organization through integrations, improving productivity (Haspeslagh and Jemison, 1991). As Handfield (1993) notes, supply integration typically means “obtaining frequent deliveries in small lots, using single or dual sources of supply, evaluating alternative sources on the basis of quality and delivery instead of price, establishing long term contracts with suppliers, reducing buffer inventories and eliminating formal paperwork.” However, integrated supply chains not only reduce costs, but also acquire partners’ technological resources and improve technology-related capacity. Unfortunately, the findings do not support H₇ and contrarily, encourage H₈. The interaction of network members can bring a shift in the business pattern and technology process.

REFERENCES

- [1] Andersson, U., Forsgren, M., and Holm, U. (2002), ‘The strategic impact of external networks: subsidiary performance and competence development in the multinational corporation’, *Strategic Management Journal*, 23(11), pp.979-996.
- [2] Dacin, T.M., Ventresca, M.J., and Beal, B.D. (1999), “The embeddedness of organizations: dialogue and direction”, *Journal of Management*, 25(3), pp.317-356.
- [3] Kogut, B. and Zander, U. (1992), ‘Knowledge of the firm, combinative capabilities and the replication of technology’, *Organization Science*, 3(3), pp.383-397.
- [4] Lee, C.-W.(2007), ‘The innovation and success of consumer electronics using new product development process’, *International Journal of Innovation and Learning*, 4(6), pp.587-611.
- [5] Marcus, A.A. and Weber, M.J. (1989), ‘Externally induced innovation’, In A.H. Van de Ven, H.L. Angle, and M.S. Poole (Eds.), *Research on the Management of Innovation: The Minnesota Studies*, Harper and Row, New York, pp.537-559.
- [6] Menanteau, P., Finon, D. and Lamy, M.-L. (2003), “Prices versus quantities: choosing policies for promoting the development of renewable energy”, *Energy Policy*, 31, pp.799-872.
- [7] Sarkis, J. (1995), ‘Supply chain management and environmentally conscious design and manufacturing’, *International Journal of Environmentally Conscious Design and Manufacturing*, 4(2), pp.43-52.

A complete form of references is available on request.

A TAGUCHI LOSS APPROACH TO SELECTION OF COLLECTION CENTERS FOR REVERSE SUPPLY CHAIN DESIGN

Kishore K. Pochampally, Department of QSOM, School of Business
Southern New Hampshire University, 2500 N. River Road, Manchester, NH 03106, USA,
(603) 629-4685, k.pochampally@snhu.edu

Surendra M. Gupta, Laboratory for Responsible Manufacturing, 334 SN, Department of MIE
Northeastern University, 360 Huntington Avenue, Boston, MA 02115, USA
(617)-373-4846, gupta@neu.edu

ABSTRACT

Many a supply chain today is involved in reprocessing of used products, which requires a series of activities that are performed by multiple parties, viz., collection centers, recovery facilities, etc. Such a supply chain is called reverse supply chain. In this paper, we propose a three-phase approach for the selection of efficient collection centers in a region where a reverse supply chain is to be designed. The first phase identifies the selection criteria, the second phase uses the eigen vector method to give relative weights to the selection criteria, and the third phase uses the Taguchi loss function to select efficient collection centers.

INTRODUCTION

Traditionally, a supply chain consists of all stages involved, directly or indirectly, in fulfilling a customer desire. But, today, many supply chains are no longer limited to fulfilling a customer desire. The reason being that the growing desire of customers to acquire the latest technology, along with the rapid technological development of new products, has led to a new environmental problem: “waste”, consisting of products that are discarded after their useful lives and the products that are discarded prematurely (hereafter, we refer to the “waste” as used products). Reprocessing of used products is essential for saving natural resources, saving energy, saving clean air and water, saving landfill space, and saving money. Besides these motivators, an important driver for companies to engage in reprocessing is the enforcement of environmental regulations by local governments [4].

Reprocessing of used products requires a series of activities (collection, disassembly, recycling, remanufacturing, disposal, etc.) that are performed by multiple parties, viz., collection centers, recovery facilities, etc., known together as a reverse supply chain. In the past decade, there has been an explosive growth of reverse supply chains, both in scope and scale.

Reverse supply chain design is a relatively new area of research, and a few case studies (with application of quantitative models) have been reported in the literature about the same. See Fleischmann’s book [1] for a detailed survey of those case studies.

This paper addresses the selection of efficient collection centers in a region where a reverse supply chain is to be designed.

SELECTION OF COLLECTION CENTERS

We propose a three-phase approach for selection of efficient collection centers in a region where a reverse supply chain is to be designed. The first phase identifies the selection criteria, the second phase uses the eigen vector method [5] to give relative weights to the selection criteria, and the third phase uses the Taguchi loss function [3] to identify efficient collection centers.

Phase – I (Identification of Selection Criteria)

We identify the following criteria for selection of collection centers:

- Process Sigma value (n) (the higher the better)
- Distance from residential area (DH) (lower distance implies higher collection)
- Distance from roads (DR) (lower distance implies higher collection)
- Simplicity of the collection process (SP) (simpler process implies higher collection)
- Per capita income of the people in the residential area (PI) (the higher the per capita income, the higher the left-over value of used products)
- Space cost (SC) (the lower the better)
- Labor cost (LC) (the lower the better)
- Incentives from local government (IG) (higher incentives from local government imply higher incentives to consumers, and hence higher collection)

The Process Sigma value (n) for a collection center represents the center's quality that could be a function of factors such as functionality of used products, efficiency of collection process, efficiency of delivery (to recovery facilities) process, and effectiveness of customer service.

Phase – II (Relative Weights of Selection Criteria)

The eigen vector method [5] begins with pair-wise judgments of importance between independent criteria with respect to the scale shown in Table 1. The normalized eigen vector of the resulting matrix of comparative importance values gives the relative weights assigned to the criteria.

Table 1. Scale for Pair-wise Judgments

Comparative Importance	Definition
1	Equally important
3	Moderately more important
5	Strongly important
7	Very strongly more important
9	Extremely more important
2, 4, 6, 8	Intermediate judgment values

The degrees of consistency of the pair-wise judgments are measured using an index called the Consistency Ratio (CR). Perfect consistency implies a value of zero for CR . However, perfect consistency cannot be demanded since, as human beings, we are often biased and inconsistent in our subjective judgments. Therefore, it is considered acceptable if CR is less than or equal to 0.1. For CR values greater than 0.1, the pair-wise judgments must be revised before the weights of criteria are computed. CR is computed using the formula:

$$CR = \frac{(\lambda_{\max} - f)}{(f - 1)I} \quad (1)$$

where λ_{\max} is the principal eigen value of the matrix of comparative importance values; f is the number of rows (or columns) in the matrix; I is the Random Index for each f value that is greater than or equal to one. Table 2 shows various I values for f values ranging from 1 to 10.

Table 2. Random Index Value for Each f Value

f	1	2	3	4	5	6	7	8	9	10
I	0	0	0.58	0.90	1.12	1.24	0.32	1.41	1.45	1.49

We consider a numerical example with the pair-wise comparison matrix shown in Table 3, for our selection criteria. For example (see Table 3), per-capita income of the people in the residential area (PI) is given 5 times more importance than the Process Sigma (n) value and 3 times more importance than the simplicity of the collection process (SP).

Table 3. Pair-wise Comparison Matrix

Criteria	n	DH	DR	SP	PI	SC	LC	IG
n	1	1	1	3	1/5	1	1	1
DH	1	1	1	1	1	1	1	1
DR	1	1	1	1	1	1	1	1
SP	1/3	1	1	1	1/3	0.2	1	1
PI	5	1	1	3	1	1	1	1
SC	1	1	1	5	1	1	1	0.2
LC	1	1	1	1	1	1	1	1
IG	1	1	1	1	1	5	1	1

Table 4 shows the relative weights of the respective criteria for selection. These weights are the elements of the normalized eigen vector of pair-wise comparison matrix shown in Table 3.

Table 4. Relative Weights of Selection Criteria

Criteria for selection	Relative weights
n	0.11
DH	0.11
DR	0.11
SP	0.07
PI	0.18
SC	0.15
LC	0.11
IG	0.16

Phase – III (Selection of Efficient Collection Centers)

In traditional systems, the product is accepted if the product measurement falls within the specification limits. Otherwise, the product is rejected. The quality losses occur only when the product deviates beyond the specification limits, thereby becoming unacceptable. These costs tend to be constant and relate to the costs of bringing the product back into the specification range. Taguchi suggests a more narrow view of characteristic acceptability by indicating that any deviation from a characteristic’s target value results in a loss. If a characteristic measurement is the same as the target value, the loss is zero.

Otherwise, the loss can be measured using a quadratic function, after which actions are taken to reduce systematically the variation from the target value. There are three types of Taguchi loss functions: “target is best”, “smaller is better”, and “larger is better”.

If $L(y)$ is the loss associated with a particular value of characteristic y , m is the target value of the specification, and k is the loss coefficient whose value is constant depending on the cost at the specification limits and width of the specification, for the “target is best” type,

$$L(y) = k(y - m)^2 \quad (2)$$

For the “smaller is better” type,

$$L(y) = k(y)^2 \quad (3)$$

For the “larger is better” type,

$$L(y) = k/(y)^2 \quad (4)$$

We now calculate the k value for each of our criteria for selection of collection centers.

Process Sigma (n) value

The loss function that applies to this criterion is “larger is better” (Equation 4). If the decision-maker considers 100% loss for n value less than or equal to 4, then the value of k is 1600%. Then, for a candidate collection center, if $n = 5$, $L(y) = 1600/(5)^2 = 64\%$. That means, with respect to quality (i.e., n value), the collection center is 36% short of the worst performance level (which in this case is “ $n \leq 4$ ”).

Distance from residential area (DH)

We consider DH as the distance of the collection center from the center of gravity [2] of all the residential areas around the center. The loss function that applies to this criterion is “smaller is better” (Equation 3). If the decision-maker considers 100% loss for DH value more than or equal to 4 miles, then the value of k is 6.25%. Then, for a candidate collection center, if $DH = 3$, $L(y) = 6.25 * (3)^2 = 56.25\%$. That means, with respect to the distance from the residential area, the collection center is 43.75% short of the worst case scenario (which in this case is “ $DH \geq 4$ miles”).

Distance from roads (DR)

We consider DR as the average distance of all the roads in the region, from the collection center of interest. The loss function that applies to this criterion is “smaller is better” (Equation 3). If the decision-maker considers 100% loss for DR value more than or equal to 5 miles, then the value of k is 4%. Then, for a candidate collection center, if $DR = 4$, $L(y) = 4 * (4)^2 = 64\%$. That means, with respect to the distance from the roads, the collection center is 36% short of the worst case scenario (which in this case is “ $DR \geq 5$ miles”).

Simplicity of collection process (SP)

Since this is a subjective criterion, the decision-maker can obtain ratings (for example, on a 1 – 10 scale, where 1 is the worst and 10 is the best) of the candidate collection centers, from experts in the field of reverse supply chain. The loss function that applies to this criterion is “larger is better” (Equation 4). If the decision-maker considers 100% loss for SP value less than or equal to 5, then the value of k is 2500%. Then, for a candidate collection center, if $SP = 7$, $L(y) = 2500/(7)^2 = 51.02\%$. That means, with respect to the simplicity of the collection process, the collection center is about 49% short of the worst case scenario (which in this case is “ $SP \leq 5$ ”).

Per capita income (PI)

The loss function that applies to this criterion is “larger is better” (Equation 4). If the decision-maker considers 100% loss for PI value less than or equal to \$30,000 per year, then the value of k is 90,000,000,000%. Then, for a candidate collection center, if $PI = \$50,000$ per year, $L(y) = 90,000,000,000 / (50,000)^2 = 36\%$. That means, with respect to the per-capita income of the people in the residential area, the collection center is 64% short of the worst case scenario (which in this case is “ $PI \leq \$30,000$ per year”).

Space cost (SC)

The loss function that applies to this criterion is “smaller is better” (Equation 3). If the decision-maker considers 100% loss for SC value more than or equal to \$1000 per day, then the value of k is 0.0001%. Then, for a candidate collection center, if $SC = \$800$ per day, $L(y) = 0.0001 * (800)^2 = 64\%$. That means, with respect to the space cost, the collection center is 36% short of the worst case scenario (which in this case is “ $SC \geq \$1000$ per day”).

Labor cost (LC)

The loss function that applies to this criterion is “smaller is better” (Equation 3). If the decision-maker considers 100% loss for LC value more than or equal to \$15 per hour, then the value of k is 0.44%. Then, for a candidate collection center, if $LC = \$10$ per hour, $L(y) = 0.44 * (10)^2 = 44\%$. That means, with respect to the labor cost, the collection center is 56% short of the worst case scenario (which in this case is “ $LC \geq \$15$ per hour”).

Incentives from local government (IG)

Since this is a subjective criterion, the decision-maker can obtain ratings (for example, on a 1 – 10 scale, where 1 is the worst and 10 is the best) of the candidate collection centers, from experts in the field of reverse supply chain. The loss function that applies to this criterion is “larger is better” (Equation 4). If the decision-maker considers 100% loss for IG value less than or equal to 7, then the value of k is 4900%. Then, for a candidate collection center, if $IG = 9$, $L(y) = 4900 / (9)^2 = 60.49\%$. That means, with respect to the incentives from the local government, the collection center is about 39.5% short of the worst case scenario (which in this case is “ $IG \leq 7$ ”).

We consider four candidate collection centers in our numerical example: C1, C2, C3, and C4. Table 5 presents the $L(y)$ for our criteria for each of the collection centers. For example, the Taguchi loss of collection center C2 with respect to the per-capita income of the people in the residential area is 52% (i.e., 48% short of the worst case scenario).

Table 5. $L(y)$ Values (%) of Collection Centers

Criteria	C1	C2	C3	C4
n (0.11)	35	54	40	63
DH (0.11)	25	15	37	55
DR (0.11)	32	10	9	90
SP (0.07)	100	75	64	50
PI (0.18)	65	52	40	50
SC (0.15)	20	10	5	100
LC (0.11)	15	18	20	67
IG (0.16)	78	64	36	41

The weighted-loss of each collection center j is calculated by using the following equation, and is presented in Table 6.

$$\text{Weighted-loss of collection center } j = \sum_i W_i L_{ij} \quad (5)$$

where W_i is the weight of criterion i (see Table 4), and L_{ij} is the Taguchi loss (see Table 5) of collection center j for with respect to criterion i . For example, the weighted-loss of C4 (see Tables 4, 5, and 6) is $0.11*63 + 0.11*55 + \dots + 0.11*67 + 0.16*41 = 64.31$.

Table 6. Weighted-Losses of Collection Centers

Collection Center	Weighted Loss
C1	45.95
C2	37.02
C3	29.85
C4	64.31

The decision-maker will select C3 because it has the lowest weighted-loss.

REFERENCES

- [1] Fleischmann, M., *Quantitative Models for Reverse Logistics: Lecture Notes in Economics and Mathematical Systems*, Springer-Verlag, Berlin, Heidelberg, 2001.
- [2] Krajewski, L. J., Ritzman, L. P. and Malhotra, M. K., *Operations Management: Processes and Value Chains, Eighth Edition*, Pearson Prentice Hall, 2006.
- [3] Pi, W. and Low, C., “Supplier evaluation and selection via Taguchi loss functions and an AHP”, *International Journal of Advanced Manufacturing Technology*, Vol. 27, pp. 625-630, 2006.
- [4] Pochampally, K. K., Nukala, S. and Gupta, S. M., *Strategic Planning Models for Reverse and Closed-loop Supply Chains*, CRC Press, Boca Rotan, Florida, 2008.
- [5] Saaty, T. L., “How to make a decision: The Analytic Hierarchy Process”, *European Journal of Operational Research*, Vol. 48, No. 1, pp. 9-26, 1990.

SELECTION OF COLLECTION CENTERS FOR REVERSE LOGISTIC NETWORKS

Onder Ondemir, Northeastern University, Boston, MA 02115, (857)-204-6090, ondemir.o@neu.edu
Surendra M. Gupta*, Northeastern University, Boston, MA 02115, (617)-373-4846, gupta@neu.edu
(*corresponding author)

ABSTRACT

In this paper, we address the problem of locating collection centers for a company that collects used products from product holders. Each product holder has an inherent willingness to return the product, and makes the decision on the basis of financial incentive offered by the company. We present a mixed-integer nonlinear facility location-allocation model to find the optimal locations, the optimal incentive values for different product types and the optimal vehicle type. Because the problem is NP-complete, a heuristic method is proposed to solve medium and large-size instances. We experiment with location dependent transportation costs which reflect the road conditions and other geographical factors. We conclude that the location dependent transportation cost approach is inferior to fixed transportation cost approach resulting in a higher loss of profit when the variation of the cost with respect to location is relatively high.

INTRODUCTION

The alarming increase in the use of virgin resources to produce new products has threatened the environment. The governments and consumers are quite aware of this phenomenon which has led to several legislations and given birth to the concept of environment consciousness. The goal of environmentally conscious manufacturing is to produce goods with minimal (or zero) harm to the environment. The concept of environment consciousness, however, now extends beyond manufacturing and includes the responsibility to manage the end-of-life products by the manufacturer [3]. The process involved in collecting, transporting and managing the EOL products is known as reverse logistics.

Reverse logistics and remanufacturing operations are more complex than forward logistics and traditional manufacturing operations. A lot of complexity stems from high degrees of uncertainties in the quality and quantity of products [2]. In reverse flow, the type of the recovery process is determined by the condition (quality) of returned products [1]. Just as the quality of returned products is unpredictable so is the quantity of returns as it may be dependent on the incentive offered to the last customer.

The type of the product recovery is dependent on the condition of a returned product. The possibilities include repairing, refurbishing, remanufacturing, cannibalizing, and recycling [4].

Used product acquisition or collection is the first step of product recovery that triggers other activities of the recovery system. As was mentioned before, there are considerable uncertainties in quantity, timing, and quality of returned products. Manufacturers can influence the quantity of returns by using buy-back campaigns and incentives to product holders. Clearly, the amount of incentive offered by the company (also called unit acquisition price) influences the quality level of used product returns. Accepting all end-of-use products in the waste stream is not a viable

strategy for most companies since a high percentage of these will have a poor quality, and hence will not be recoverable. As a result, a used product acquisition strategy, by offering an appropriate incentive, is crucial for a company engaged in product recovery [1].

In this paper, we present a model to find the optimal locations of a predetermined number of collection centers and the optimal incentive that should be offered by the company to product holders based on the quality condition of their used items. In this model we consider a pick-up scenario in which the company collects used products from the premises of the product holders and takes responsibility of all the collection-related costs. It is assumed that the willingness of the product holders to return is effected by the amount of the financial incentive offered. Thus, a higher incentive reduces the unit cost savings from a return, yet increases the product holders' willingness to return their used products.

MODEL

Assumptions

We assume that a pick-up strategy is in place according to which vehicles with limited capacity are dispatched from the collection centers to the locations of product holders to transport the returns. There are three vehicle alternatives (small, medium and large size) to be considered. There are no capacity constraints on collection centers. The cost per unit distance traveled between a customer zone and a candidate site is not fixed for all zone-site pairs. We also assume that the financial incentive offered by the company determines the willingness of product holders to return their used products.

Nomenclature

Parameters:

m	number of candidate sites,
n	number of customer zones,
d_{ij}	travel distance between customer zone j and, candidate site i ,
L	number of vehicle alternatives,
c_{1l}	operating cost of a vehicle of type l ,
c_{2ij}	cost per unit distance traveled between site i and zone j ,
q_l	capacity of a vehicle of type l ,
h_j	number of product holders living in zone j ,
K	number product types,
γ_{jk}	proportion of product holders of type k in zone j ,
h_{jk}	number of product holders of type k living in zone j ($h_{jk} = \gamma_{jk}h_j$),
s_k	unit cost savings from a used product of type k ,
a_k	the maximum incentive level of product holder of type k ,
m_l	travel cost coefficient for vehicles of type l ,
p	number of collection centers to be opened.

Variables:

Π	total profit,
X_{ijk}	fraction of potential returns of type k collected in zone j and transported to the CC at site i ,

V_{ij} number of vehicles required to transport returns from zone j to the CC at site i ,
 R_k unit incentive offered (unit acquisition price) for a used product of type k ,

$Y_i = \begin{cases} 1, & \text{if a collection center (CC) is located at site } i \\ 0, & \text{otherwise} \end{cases}$,
 $t_l = \begin{cases} 1, & \text{if a vehicle of type } l \text{ is selected} \\ 0, & \text{otherwise} \end{cases}$.

Formulation

$$\max \Pi = \sum_i^m \sum_j^n \sum_k^K \frac{X_{ijk} h_{jk} R_k (s_k - R_k)}{a_k} - \sum_i^m \sum_j^n \sum_l^L (c_{1l} t_l + 2c_{2l} m_l d_{ij}) V_{ij} \quad (1)$$

s.t.

$$\sum_{i=1}^m X_{ijk} \leq 1, \quad j = 1, \dots, n; \quad k = 1, \dots, K, \quad (2)$$

$$X_{ijk} \leq Y_i, \quad i = 1, \dots, m; \quad j = 1, \dots, n; \quad k = 1, \dots, K, \quad (3)$$

$$\sum_{i=1}^m Y_i = p \quad (4)$$

$$V_{ij} \geq \frac{\sum_k X_{ijk} h_{jk} R_k (s_k - R_k)}{\sum_l t_l a_l}, \quad i = 1, \dots, m; \quad j = 1, \dots, n, \quad (5)$$

$$\sum_{i=1}^L t_i = 1 \quad (6)$$

$$R_k \leq a_k, \quad k = 1, \dots, K, \quad (7)$$

$$R_k \leq s_k, \quad k = 1, \dots, K, \quad (8)$$

$$R_k \geq 0, \quad k = 1, \dots, K, \quad (9)$$

$$X_{ijk} \geq 0, \quad i = 1, \dots, m; \quad j = 1, \dots, n; \quad k = 1, \dots, K, \quad (10)$$

$$V_{ij} \geq 0 \text{ and integer}, \quad i = 1, \dots, m; \quad j = 1, \dots, n, \quad (11)$$

$$Y_i \in \{0,1\}, \quad i = 1, \dots, m, \quad (12)$$

$$t_l \in \{0,1\}, \quad l = 1, \dots, L. \quad (13)$$

SOLUTION METHODOLOGY

The problem formulated as a mixed-integer nonlinear programming model is very difficult to solve even for small problem sizes. For this reason we developed a solution method which reduces the complexity of the problem.

The main loop of the method is based on a tabu search method performed in the space of collection center locations. For each location set prescribed by tabu search, simplex search (Nelder-Mead Simplex) is called to obtain the best incentives and the corresponding net profit (Figure 1). This main loop is run for each vehicle alternatives.

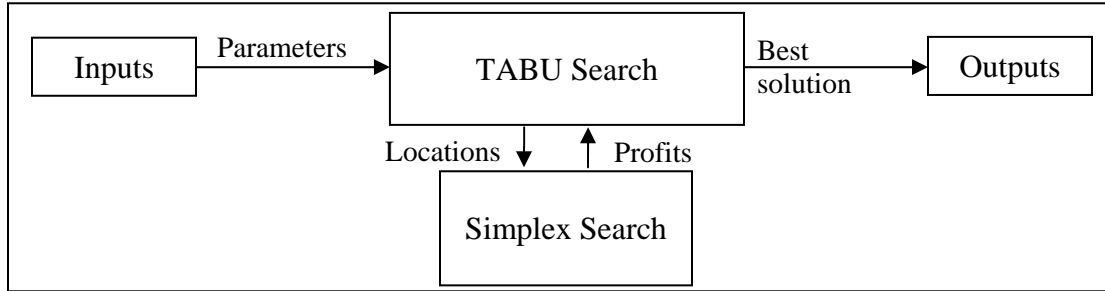


FIGURE 1. Basic flow chart of the proposed solution method.

COMPUTATIONAL RESULTS

For two distinct values of n ($n=10, 20$) and three distinct values to p ($p=2, 3, 4$), we obtain 6 instances and solve them using commercial solvers BARON (in GAMS 22.5) and LINGO 7.0, and using the algorithm described above. Proposed algorithm has been coded in MATLAB 6.1 and all experiments have been conducted on a laptop computer with 1.60 GHz Intel Celeron M processor and 512 MB ram.

Since the worst computation time of the proposed heuristic algorithm is 481.781, we set a time limit of 600 sec. in GAMS and LINGO. GAMS could find a local optimum for the smallest instance whereas LINGO could not find any feasible solution. Comparative results are shown in Table 1. For instances (20x2x2) and (20x3x2), large size vehicles, and for the rest of the instances, medium size vehicles turn out to be the optimal vehicle selection.

TABLE 1. Computational results comparisons

Instance	BARON (GAMS)			Proposed Heuristic			
	R1	R2	Objective	R1	R2	Objective	CPU time (s)
10x2x2	9.841	5.000	1,762.426	10.7621	5.7586	1,890.7	36.125
10x3x2	-	-	-	9.2174	4.4546	2,479.4	91.281
10x4x2	-	-	-	10.6520	5.8922	2,298.9	96.61
20x2x2*	-	-	-	9.2294	4.6540	4,707.1	205.671
20x3x2*	-	-	-	9.0970	4.5468	4,908.3	454.89
20x4x2	-	-	-	9.5697	4.8706	4,432.8	481.781

CONCLUSIONS

In this paper, we presented a mixed integer non-linear programming model to find the optimal locations of a predetermined number of collection centers, the optimal vehicle type and the optimal incentive that should be offered by the company to product holders based on the quality condition of their used items. In order to solve this NP-complete problem, a tabu search based algorithm was developed in MATLAB.

Experiments for 6 instances have shown that the vehicle type has a significant effect on the collection policy. We also conclude that the location dependent transportation cost approach is inferior to fixed transportation cost approach resulting in a higher loss of profit when the variation of the cost with respect to location is relatively high.

REFERENCES

- [1] Aras, N., Aksen, D., Tanugur, A. G. "Locating collection centers for incentive-dependent returns under a pick-up policy with capacitated vehicles." *European Journal of Operational Research* (in press), 2007.
- [2] Gungor, A., Gupta, S. M. "Issues in environmentally conscious manufacturing and product recovery: a survey." *Computers & Industrial Engineering*, 1999, 36(4), 811-853.
- [3] Kara, S., Rugrungruang, F., Kaebnick, H. "Simulation modelling of reverse logistics networks." *International Journal of Production Economics*, 2006, 106(1), 61-69.
- [4] Thierry, M., Salomon, M., van Nunen, J., van Wassenhove, L.N. "Strategic issues in product recovery management." *California, Management Review*, 1995, 37, 114-135.

OUTSOURCING REVERSE LOGISTICS FUNCTIONS

Sharon Ordoobadi

University of Massachusetts-Dartmouth

285 Old Westport Road, N. Dartmouth, MA. 02747

sordoobadi@umassd.edu

508-999-8767

ABSTRACT

A phased decision model is developed that starts with a strategic analysis to examine if it is strategically sound to outsource any of the reverse logistics activities. If these activities are considered to be core competencies of the organization, then no further analysis is required and they will be performed in-house. The reverse logistics activities that are identified as peripherals are contracted out to a third party. Activities that can not be identified as core competencies or peripherals by strategic analysis alone will be tested against several other significance factors. Those activities that are identified as highly significant to the success of the organization will be performed in house. While low significance activities are candidates for being outsourced. In addition, the economic evaluation phase of the model tests the financial feasibility of such actions. Thus, the result is a decision that is both strategically sound and economically plausible.

BACKGROUND AND RESEARCH MOTIVATION

In recent years companies have spent a lot of time and money on their reverse supply chain. The financial impact of reverse logistics on company's profits is evident from the huge costs allocated for these activities. The reverse logistics costs for 1999 reported as \$37 billion dollars. These costs are estimated as 4% of the total logistics costs, while logistics costs account for approximately 10% of the U.S. economy (Rogers & Tibben-Lembke, 2001). The reverse logistics costs are quite high and are increasing because of the liberal return policies that manufacturers/retailers have to offer to keep their competitive advantage.

Manufacturers choose to engage in reverse logistics and specially product recovery to reduce production cost, meet customer demand, enhance brand image, and protect after market (Toffel, 2004). In some cases manufacturers are forced to set up reverse supply chains because of environmental regulations or to meet customer requests or both (Guide & Wassenhove, 2002). Whether companies are creating reverse supply chains by choice or by force, they need to choose between performing the function in their own plant and contracting it out to a third party.

All of the above, point to the strategic importance of reverse logistics activities and in particular the impact of outsourcing such activities on competitive position of the company. Thus, the question of whether to outsource reverse logistics activities is addressed in this research. This is done by developing a model that allows the decision makers to analyze the impact of their outsourcing decisions as well as guiding them through their decision making process. The proposed model has four phases: strategic, significance, economic, and decision phase. The model starts with strategic analysis and proceeds to the next phase or ends depending on the result of the analysis. If strategic analysis determines that the reverse logistics activity is considered a core competency, then no further analysis is required and that activity is performed

in-house. Otherwise the model proceeds to the second and third phase namely significance and economic analysis. The results of the significance and economic analysis are then combined to determine the final course of action.

The rest of the paper is organized as follows: Section II provides reasons for outsourcing reverse logistics activities as well as some of the risks involved. The proposed model is introduced in section III. Finally the paper closes with conclusions in section IV.

II. BENEFITS AND RISKS OF OUTSOURCING REVERSE LOGISTICS FUNCTIONS

Several reasons for outsourcing reverse logistics activities to a third party logistics provider are cited by researchers (Wilding, 2004, Bolumole, 2001, Sink & Langely, 1997, Ferrer & Whybark, 2001):

- Competencies of the 3PL provider
- Operational flexibility/efficiency
- Avoiding large investments
- Expansion to new markets
- Labor issues
- Lack of internal expertise
- Cost advantages due to economies of scales
- Ability to focus on core activities
- Improved customer service/ Ability to provide customized service
- Handling of the non-value added activities
- Gain access to the world-class technology
- Availability of more specialized logistics expertise
- Reduction in employee base
- Reduction in operating and transaction costs
- Differentiation from competitors

The risks of outsourcing reverse logistics functions to a third party cited by some researchers (Bolumole, 2001, Sink & Langely, 1997, Earl, 1996) are listed as:

- Increased dependence on the service provider
- Access to the company's information by the 3PL provider.
- Loss of control on the outsourced activity
- Negative impact on customer service
- No reduction on time and effort spent on reverse logistics activities
- Low quality service
- Increase in customer complaints
- Loss of innovative capacity
- Loss of customer loyalty

III. THE OUTSOURCING MODEL

Reverse logistics comprises all activities required in order to convert the products no longer needed by the user to products usable again in a market. The main categories of activities are: distribution planning, inventory management, and production planning (Fleischmann et al, 1997). Distribution planning refers to physical transport of the unwanted products from end users to producers. Possible activities in the reverse distribution channel are: collection, testing, sorting, and transportation. Inventory management refers to managing the transformation of the returned products to usable products. Production planning refers to disassembly, recycling,

repair, remanufacturing, and disposal. The purpose of some of these activities (e.g. recycling) is material recovery, while others (e.g. remanufacturing) are carried out for the purpose of product recovery.

Any of the activities within the main functions described above are candidates for being outsourced to a third party service provider. Reverse logistics activities can be considered as operational or “basic logistics” such as warehousing, and strategic or “value-added logistics” such as remanufacturing (Berglund, et al, 1999). The operational activities are outsourced as long as the outsourcing option is financially justified. The value-added logistics are multi-dimensional and have business value to the company. Proper handling of these activities would increase the business value and thus, strengthen the competitive position of the company. On the other hand, mishandling them could be detrimental to the company’s profit and its competitive position. This research concentrates on strategic and value-added activities of the reverse supply chain. The remanufacturing process of the reverse logistics is chosen as the focus of this study. However, the analysis can be applied to other reverse logistics activities/processes as well. Remanufacturing refers to those activities that bring used products back to such a condition that recovered items are just as good as new (Inderfurth, 2005). The original manufacturer may decide to outsource the entire process or selected activities within this process. The four phases of the model is explained in detail in the following sections.

Strategic Analysis

Strategic analysis starts by performing a core competency check.

Core Competency Check

Core competencies refer to activities that distinguish a company from its competitors. An activity has to satisfy several criteria to be classified as a core competency (Prahalad & Hammel, 1990; Quinn & Hilmer, 1994). The following questions are used to determine if an activity is a core.

- Does the activity need highly specialized skills?
- Does the activity have a high impact on what customers perceive as the most important product attribute?
- Does the activity provide potential access to a wide variety of possible future markets?
- Is the activity difficult for competitors to duplicate?

An activity is considered a high core competency if there are at least three “yes” answers to the questions asked. An activity is considered a low core competency if there are one or two “yes” answers. If there are no “yes” answers, the activity is considered peripheral. High core competencies are performed in-house and no further analysis is needed, since these competencies offer long term competitive advantage and must be controlled, invested in, and protected. Peripheral activities will be contracted out to a third party. Outsourcing these activities is not detrimental to the company’s competitive position. However, economic analysis is performed to test the financial feasibility of such an action. If it is cheaper to contract with a third party logistics provider, then outsourcing option is both strategically and financially sound. On the other hand, if outsourcing costs exceed the costs of performing the activity in-house the decision maker can weigh the benefits of still outsourcing and investing in another area with a considerable high rate of return. Low core competencies are somewhat marginal and require further analysis to determine the appropriate policy. These marginal activities will be tested against several significance factors to determine their significance to the success of the

organization. This evaluation is possible by conducting the significance analysis. Thus, the model proceeds to the next phase, significance analysis which is detailed in the following section.

Significance Analysis

In this phase, the low core competencies are evaluated for their significance to the company. The significance refers to whether the reverse logistics activity is one of company's critical success factors. The following questions are used to check the significance of an activity (Toffel, 2004, Maltz, 1994).

- Does the reverse engineering provide better and speedier knowledge for the designers?
- Is there a possibility of improving competitive position?
- Is there a high risk of the Original Equipment Manufacturers (OEMs) becoming dependent on external product recovery firms?
- Is there a great uncertainty regarding product return?
- Are the returned products high-value items?
- Is company following differentiation strategy?
- Is size of the company large?

A positive response to any of the above questions points to the significance of an activity. An activity is identified as highly significant to the success of the organization if there are positive responses to all the questions. On the other hand, if there are not any positive response to the above questions the activity is considered to have a very low significance to the success of the company. Activities that do not fall into either of the classifications are identified as moderately significant. It seems logical to keep the highly significant activities in-house, the impact on the organization is too significant to let a third party performs these activities. Activities that have very low significance are good candidates for being outsourced to a third party. Those activities that are identified as moderately significant are somewhat marginal and might better be handled by vertical integration or hybrids. However, the economic ramifications of each of these actions should be analyzed before an action is taken. The economic analysis phase of the model will evaluate cost advantages/disadvantages of each of the actions.

Economic Analysis

Economic analysis is performed to compare the cost of performing a reverse logistics activity in-house with the cost of contracting it out to a third party. Since this research focuses on the remanufacturing function, the costs of performing the remanufacturing process in-house and contracting it out to an independent remanufacturer will be identified. These costs are then compared to conduct the economic analysis.

The costs generally associated with performing the remanufacturing function in-house are identified as: labor costs, capital costs, material costs, inventory holding costs, and overhead costs. The definition of each of these cost elements follows:

Labor costs (LC): Costs of the labor required for disassembly, inspection, reassembly, and repackaging.

Material costs (MC): Costs of obtaining materials for remanufacturing purposes.

Capital Costs (CC): Costs of all equipment, facilities, and tools that are used in remanufacturing.

Inventory holding costs (IHC): Costs of holding the returned products in storage.

Overhead costs (OVC): Refers to administrative costs, utility costs, and other costs attributed to the remanufacturing function.

Let TC (in-house) denotes the total costs of performing remanufacturing function in-house. Then $TC(\text{in-house}) = LC + MC + CC + IHC + OVC$

Of-course there are certain costs associated with outsourcing the remanufacturing function to an independent remanufacturer. These costs are categorized as: information costs, transportation costs, price (sub-contract costs), and administrative costs. The definition of these cost elements follows:

Information costs (IC): Costs related to seeking information on the potential third party service providers and locating a suitable remanufacturer.

Transportation costs (TRC): Costs of transporting returned products to the remanufacturer plant.

Sub-contracting costs (SC): The contracted costs of remanufacturing the items charged by the remanufacturer.

Administrative costs (AC): Costs of negotiating and drawing up of contracts, monitoring the remanufacturer's performance, renegotiating contracts, legal expenses, and other costs attributed to dealing with the remanufacturer.

Let TC (outsource) denotes total costs of contracting out the remanufacturing function to a third party (an independent remanufacturer). Then $TC(\text{outsource}) = IC + TRC + SC + AC$

To allow comparison between the total costs of outsourcing remanufacturing function and performing it in-house, a delta cost function is introduced. The Δcost function is defined as the difference between total costs of performing the activity in-house and contracting it out to a third party. Then $\Delta\text{cost} = TC(\text{in-house}) - TC(\text{outsource})$

If $\Delta\text{cost} > 0$ outsourcing remanufacturing function has cost advantage

If $\Delta\text{cost} < 0$ performing remanufacturing in-house has cost advantage

The result of the economic evaluation will be used in the final phase of the model, decision analysis, to determine the recommended course of action.

Decision Analysis

The results of the significance evaluation and economic analysis are combined to provide a guideline for managerial decision making. This is done by combining the rankings from significance analysis and the delta cost function. Activities that have low significance with positive delta costs or near zero delta costs will most likely be contracted out to a third party. These activities are not that important to the organization and they can be performed more economically by a third party logistics provider.

For activities that have low strategic significance but have in-house cost advantage it is recommended to perform internally if the cost advantage is very significant. However, if the savings is not that huge the company might consider the outsourcing option in order to free-up money/asset that would have otherwise been used to support the in-house remanufacturing. Performing remanufacturing function internally requires huge initial capital investment and fixed costs that can be saved if the process is outsourced. This freed-up money can be diverted to another avenue with a significant rate of return. However, to determine if such a decision is

sound a Rate of Return (R.O.R.) analysis is conducted. First, the Internal Rate of Return (IRR) for a specific investment is calculated. This rate is then compared with the Minimum Attractive Rate of Return (MARR), the return that a company requires to justify its investment. If IRR is greater than MARR, then it might be wise to outsource that activity and invest the freed-up money. Otherwise, the investment is not justified and it is recommended to perform such an activity internally.

Activities that have high significance with negative delta costs will be kept in-house. They are strategically significant to the company and in-house performance has cost advantage. Highly significant activities that can be performed cheaper by a third party should be treated cautiously. Even though it is economically justified to outsource these activities, it is vital that the company keeps tight control over such significant activities. Thus, it is recommended to do “smart” sourcing, that is outsourcing with close relationship with the third party to guarantee the strategic control remains with the organization. In other words, these types of activities require relational agreement with the third party rather than just contractual agreement.

Decisions regarding activities that are moderately significant are more financially driven. If outsourcing an activity has cost advantage then that activity will be contracted out to a third party. However, the manufacturer should keep some degree of control over the performance of the third party logistics provider. The activity is moderately important to the success of the organization and loss of control by the manufacturer could be damaging to the company. Thus, it is recommended to form a bilateral contract with the third party rather than just contractual or arm’s length agreement. If a moderately significant activity can be performed more economically in-house it is recommended to perform internally. The importance to the organization is moderate, and performing it in-house has cost advantage.

IV. CONCLUSIONS

The decision regarding outsourcing of reverse logistics activities is considered a strategic decision. Thus, first the benefits and risks of outsourcing these activities to a third party service provider were identified. The focus of the research however is on the development of a decision model to help manufacturers make more informed decisions regarding their reverse logistics policies. The proposed model starts with strategic analysis to determine if it is strategically sound to outsource the reverse logistics activities. If these activities are determined to be core competencies of the organization, then no further analysis is required and the activities are performed in-house.

Reverse logistics activities that are identified as peripherals will be contracted out to a third party. Activities that cannot be identified as core competency or peripheral by strategic analysis alone will be tested against several other significant factors. Reverse logistics activities that are identified as highly significant to the success of the organization will be performed in house. Activities that have low significance are candidates for being contracted out to a third party. The economic evaluation phase of the model tests the financial feasibility of such actions. Thus, the result is a decision that is both strategically sound and economically plausible.

(References are available upon request from Sharon Ordoobadi)

RELATIONSHIP BETWEEN FIRM PERFORMANCES AND PROFITABILITY

S. Bruce Han, Girard School of Business and International Commerce, Merrimack College,
North Andover, MA 01845, (978) 837-5319, Bruce.Han@merrimack.edu

INTRODUCTION

In this study several research questions will be empirically investigated. The research objective is to conduct an empirical study based on real data to better understand the relationship between ISO 9001 certification efforts and firm performances, as well as their effects on profitability. Thus, a model is developed and proposed based on the literature (see Figure 1), and several hypotheses are posed and tested.

Nine hypotheses to be tested are as follows:

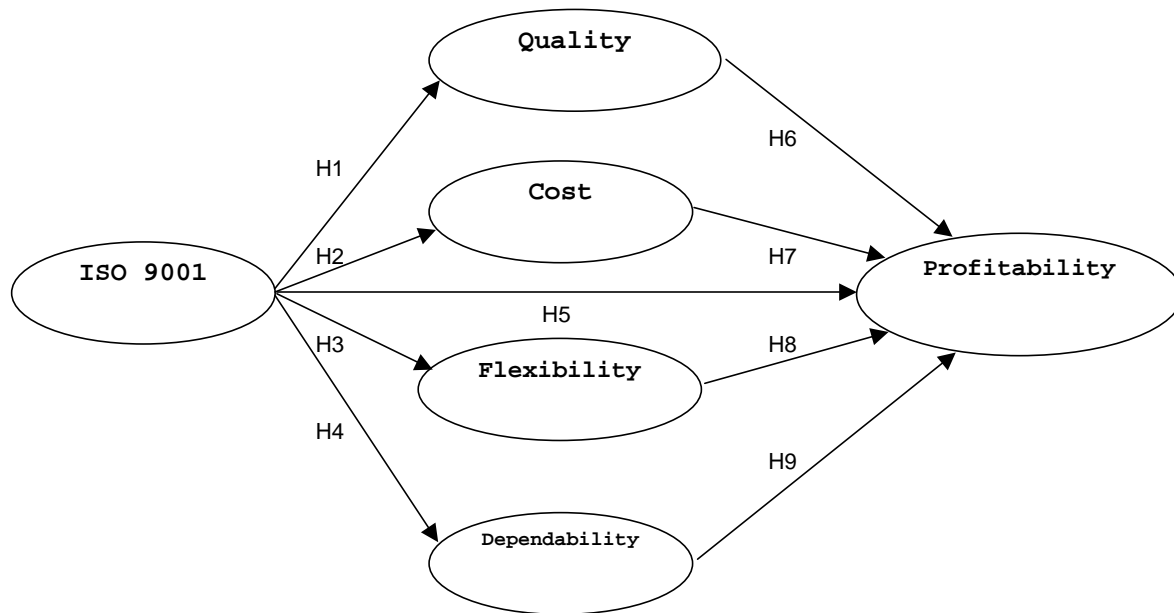
- H1: ISO 9001 registration efforts have positive effect on quality.
- H2: ISO 9001 registration efforts have positive effect on cost reduction.
- H3: ISO 9001 registration efforts have positive effect on flexibility.
- H4: ISO 9001 registration efforts have positive effect on dependability.
- H5: ISO 9001 registration efforts have positive effect on profitability.
- H6: Quality has positive effect on profitability.
- H7: Cost reduction has positive effect on profitability.
- H8: Flexibility has positive effect on profitability.
- H9: Dependability has positive effect on profitability.

METHODS

The sample of this study consists of 441 ISO 9001 registered manufacturing companies operating in the United States. The survey instrument that was mailed to ISO 9001 champions of these firms utilized a five-point Likert scale with 1 representing the low end of the scale and 5 representing the high end. A total of 2130 firms were surveyed. Of these, 1600 belonged to the electronic/electrical industry and 530 belong to the chemical industry. These two industries were selected because electronic/electrical industry was the first industry that embraced ISO 9000 and is best represented in terms of number of companies registered in the United States. The chemical industry was chosen because it is sufficiently different from electronic/electrical industry and is also well represented in terms of number of companies registered in the United States. It is also one of the major industries often mentioned in empirical studies done in the area of quality management. These include all of the firms with their ISO 9001 champions identified.

ISO 9001 champions are determined as the most appropriate respondents since they are most familiar with their firm's quality practices and its impact on firm's performances and profitability. Of the 2130 mailed surveys, 187 were returned as undeliverable. The most frequently mentioned reason for undeliverable surveys was that the addressee was no longer (at this address) with the company. Within two months of mailing, 445 completed surveys were returned giving a 23% response rate. Upon review, the responses of 4 surveys were deemed to be unusable due to missing or invalid information resulting in 441 usable responses.

Figure 1 The theoretical model



RESULTS

Table 1 presents the reliability and convergent validity of the constructs. Also all six constructs explain more than 50% of the variance indicating the discriminant validity. Table 2 shows that all constructs are positively correlated. Seven hypotheses (H1, H2, H3, H4, H6, H7, and H8) are supported at the .01 significance level. Also, two hypotheses (H5 and H9) are not supported (see Table 3 and Figure 2).

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 performances. The results of the analysis also support the belief that ISO 9001 contributes to competitiveness, which in turn helps to gain greater profitability.

Table 1 Reliability and convergent validity of constructs

Construct and Indicators	Standardized Loading	Reliability	Variance Extracted
ISO 9001		.84	.52
Quality System	.71		
Document and Data Control	.73		
Process Control	.70		
Corrective and Preventive Action	.74		
Quality Records	.75		
Internal Quality Auditing	.76		
Training	.66		
Quality		.90	.77
Conformance to Design Specification	.77		
Product Performance	.92		
Product Durability	.90		
Product Reliability	.91		
Cost		.89	.70
Unit Production Cost	.85		
Inventory Level	.82		
Capacity Utilization	.87		
Productivity	.86		
Waste	.80		
Flexibility		.83	.64
Customization	.78		
Rapid Capacity Adjustment	.88		
Product Design Time	.77		
Set Up Time	.76		
Dependability		.94	.89
Delivery Speed	.94		
Delivery Reliability	.95		
On-Time Delivery	.94		
Profitability		.96	.89
Net Income	.94		
Revenue	.97		
Financial Performance	.95		
Total Sales	.91		

Note: All t tests were significant at $p < .01$

Table 2 Correlation matrix – constructs

	<u>ISO</u>	<u>Quality</u>	<u>Cost</u>	<u>Flexibility</u>	<u>Dependability</u>	<u>Profit.</u>
ISO	1.000					
Quality	0.601	1.000				
Cost	0.564	0.339	1.000			
Flexibility	0.573	0.344	0.323	1.000		
Dependability	0.566	0.340	0.319	0.324	1.000	
Profitability	0.538	0.518	0.654	0.468	0.283	1.000

Note: All were significant at $p < .01$

Table 3 Results of the hypotheses test

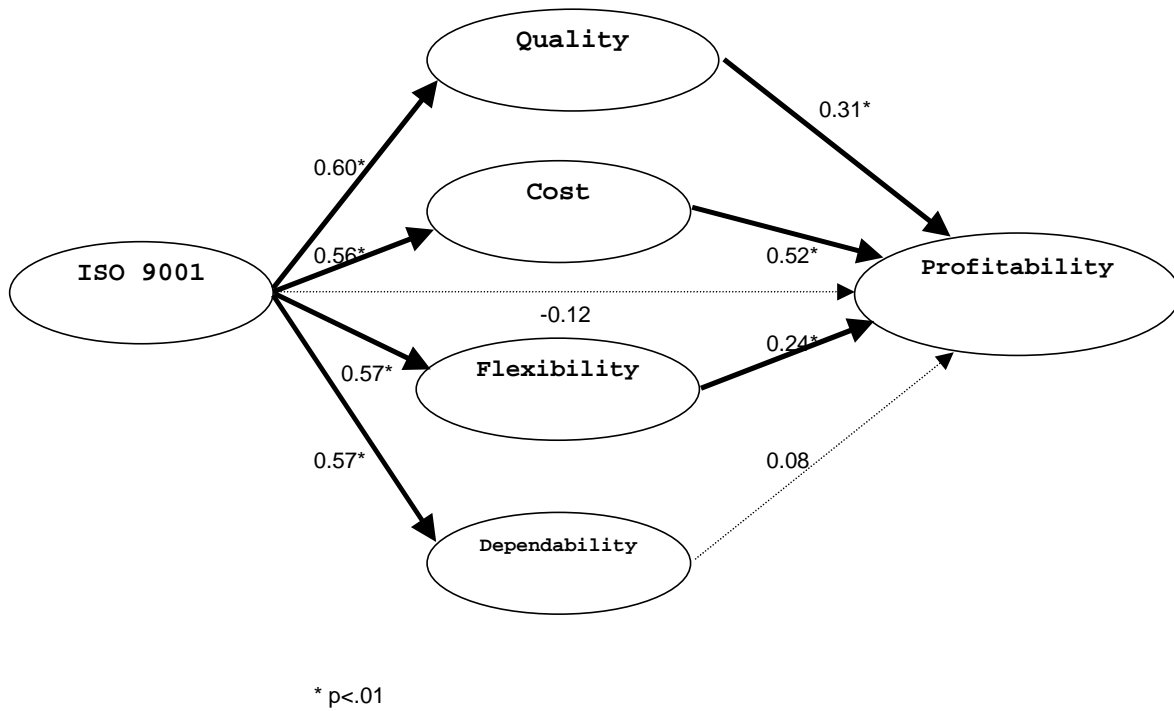
Hypothesis	From	To	Standardized Coefficient	t	Hypothesis supported?
H1	ISO	Quality	0.60	8.87*	yes
H2	ISO	Cost	0.56	9.15*	yes
H3	ISO	Flexibility	0.57	8.73*	yes
H4	ISO	Dependability	0.57	9.69*	yes
H5	ISO	Profitability	-0.12	-1.51	no
H6	Quality	Profitability	0.31	5.87*	yes
H7	Cost	Profitability	0.52	9.73*	yes
H8	Flexibility	Profitability	0.24	4.72*	yes
H9	Dependability	Profitability	0.08	1.83	no

* $p < .01$

Table 4 Measures of fit for the structural model

Fit measure	Recommended value	Structural model
Chi-Square		1314.52
d.f.		315
p-value		0.000
Chi-Square/d.f.		4.17
CFI	> 0.95	0.97
TLI	> 0.95	0.96
RMSEA	< 0.1	0.09

Figure 2 The structural model



CONCLUSIONS

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. For example, this study supports that there is a significant, positive relationship between ISO 9001 registration efforts and firm performances. Furthermore, this study supports that ISO 9001 registration efforts do not have a direct, positive relationship with profitability. Instead, this study suggests that ISO 9001 registration efforts help to reduce cost, improve quality, and increase flexibility, which in turn enhances profitability. Lastly, the findings were relevant and important to practicing managers as well as quality/operations management researchers.

REFERENCES

- Anderson, S. W., Daly, J. D. and M. F. Johnson. (1999). Why firms seek ISO 9000 certification: Regulatory compliance or competitive advantage?. *Production and Operations Management*. 8.1: 28-43.
- Douglas, A., Coleman, S. and R. Oddy. (2003). The case for ISO9000. *The TQM Magazine*. 15.5: 316-324.
- Foster, S. T. (2007). *Managing quality : Integrating the supply chain*, 3rd edition. Upper Saddle River, New Jersey: Pearson Education, Inc.
- Han, S. B. and S. K. Chen. (2007). Effects of ISO 9000 on customer satisfaction. *International Journal of Productivity and Quality Management*. 2.3: 208-220.
- Han, S. B., Chen, S. K. and M. Ebrahimpour. (2007). The Impact of ISO 9000 on TQM and Business performance. *Journal of Business and Economic Studies*. 13.2: 1-23.
- Lakhal, L, Pasin, F. and M. Liman. (2006). Quality management practices and their impact on performance. *International Journal of Quality and Reliability Management*. 23.6: 625-646.
- Naveh, E. and A. Marcus. (2005). Achieving competitive advantage through implementing a replicable management standard: Installing and using ISO 9000. *Journal of Operations Management*. 24.1: 1-26.
- Poksinska, B., Eklund, J and J. J. Dahlgaard. (2006). ISO 9001:2000 in small organisations: Lost opportunities, benefits and influencing factors. *International Journal of Quality & Reliability Management*. 23.5: 490-512.

CHALLENGES AND FRUSTRATIONS OF IMPLEMENTING LEAN MANUFACTURING

Khim Ling Sim, Gabelli School of Business, Roger Williams University, (401) 254-3101, ksim@rwu.edu
John W. Rogers, School of Bus Adm, American International College, (413) 205-3230, jrogers@acad.aic.edu
John McQuilkin, Gabelli School of Business, Roger Williams University, (401) 254-3776, jmcquilkin@rwu.edu

ABSTRACT

This study was conducted to help us understand why implementing continuous improvement strategies can be difficult at times. It also addresses the problem of resistance to change within even those firms whose CEO is most fully committed to implementation of continuous improvement programs. The study is based on a Fortune 500 manufacturing plant located in the United States. A survey was distributed to both salaried and unionized hourly employees. Results from the survey show that the problem lies primarily with an aging and high seniority hourly workforce and a lack of committed leadership at this research site.

INTRODUCTION

Producing more with less - is often the key to survive in this competitive environment. Changing production methods from mass-production with high inventory to a leaner operation with low inventory has become an essential practice for successful manufacturers, including General Electric, United Technology Corporation, Honeywell, among others [1]. The message to employees is clear - if goods cannot be produced defect free, on time, and at a reasonable price, they will be manufactured elsewhere [2].

The above story is well known and often told in corporate America. What is less often discussed is the depth of resistance to change that impedes the implementation of these techniques. Streamlining an established industrial firm can bring management into conflict with many subtle but deeply entrenched attitudes that resist change and the fundamental shift toward greater efficiency that these "lean" techniques promise. This is true even in the case of firms whose CEO is fully committed to quality and ready to spend significant corporate resources in explaining these programs to all levels of employees and training them thoroughly in their implementation [3].

This research project addresses the problem of resistance to change within even those firms whose CEO is most fully committed to implementation of Continuous improvement programs. Collaborators on this research project worked closely in the implementation of continuous improvement initiatives within a major corporation and struggled as they saw events fall short of their intended goals. Many implemented programs fell back to their original costly and chaotic mode. To develop an understanding of these complex organizational dynamics, a survey was conducted to gather data directly from the management and workers on the shop floor.

Total Quality Systems and Lean Production

Improving quality has been widely recognized as a competitive advantage in the global economy [4]. Brah et al. (2002) supported the proposition that Total Quality implementation correlates with quality conformance [5]. Thus, Total Quality Systems aim at improving quality. Lean Production or lean management, on the other hand, originated at Toyota in Japan by Japanese engineers Taiichi Ohno and Shigeo Shingo and has been implemented by many major US firms, including Danaher Corporation

and Harley-Davidson [6]. The driving force behind the development of lean management was the elimination of waste, especially in Japan, a country with few natural resources.

Both total quality systems and lean production or lean management have now evolved into comprehensive management systems. Their effective implementation involves cultural changes in organizations, new approaches to product and to serving customers, and a high degree of training and education of employees, from upper management to the shop floor. Thus, both systems have many overlapping features, such as an emphasis on customer satisfaction, high quality, comprehensive employee training, empowerment, management commitment, and communication, among others [6].

Litton (2001) has shown that effective TQM processes can result in marked improvements in both product and service quality which drive increased customer satisfaction and organizational profitability [7]. Antony et al. (2002) suggested that total quality implementation will result in employee involvement, improved communication, increased productivity, improved quality and less rework, as well as improved customer satisfaction [8]. Finally, Samat et al. (2006) report that successful implementation of total quality systems demands full commitment from various parties in the organization and will require some changes and restructuring that involves a large amount of time [9].

METHODOLOGY

This study is based on a manufacturing plant located in the Eastern United States. It is part of a Fortune 500 Company. For nearly a decade, this company has gone through a transformation from a mass production facility to a Just in Time lean manufacturer. Changes have been met with much resistance from the salary and unionized hourly workforce. A random sampling of 100 employees was requested to participate in a survey study. This sample size represents about 30% of the total employee who work within the first shift. The survey was broken into two main sections. It started with four questions to obtain demographic data. The second part of the questionnaire included 20 statements, which was divided into 5 categories. For each area of interest, four questions were included. Responses to the questions range on a Likert scale from 1-5, i.e., strongly agree (1) to strongly disagree (5).

RESULTS

Descriptive statistics based on the demographic information were presented first followed by detailed findings from the survey.

Findings

Table 1a and 1b show the demographic Information of the respondents.

Table 1a – Descriptive Statistics, By Job Category and Gender, Frequency

	Hourly vs. Salary		Gender		
	Frequency	Percent		Frequency	Percent
Hourly	45	45.5	Male	68	68.7
Salary	54	54.5	Female	31	31.3
Total	99	100.0		99	100.0

**Table 1b – Descriptive Statistics
By Education and Year of Service, Frequency**

Education			Year of Service		
	Frequency	Percent		Frequency	Percent
High School	48	48.5	0 - 5 yrs	17	17.2
Associate	16	16.2	6 -10 yrs	5	5.1
Bachelor	26	26.3	11- 20 yrs	6	6.1
Post Graduate	9	9.1	21 – 30 yrs	41	41.4
			30 yrs +	30	30.3
Total	99	100.0	Total	99	100.0

Next, results of Principal Component analysis with varimax rotation were tabulated in Table 2a. The twenty items loaded unto 4 factors all with Eigenvalue greater than 1 with total variance explained of 69.5%. Table 2b shows the mean responses from the four resulting factors, standardized alphas were also reported. As shown, the alphas range from a low of 0.75 to a high of 0.85. Table 2c presents results of Pearson’s correlations. On average, hourly employees have longer tenure in this plant. In addition, year of service is associated with unfavorable ratings (i.e., disagree or strongly disagree) in “practice and Operations”, “Training & Development”, and “Job Security”. Finally, note that unfavorable ratings in “Training and Development” and “Job Security” are highly associated with hourly employees.

**Table 2a – Principal Component Analysis
Rotated Component Matrix^a**

	Component			
	Practice & Operations	Job Security	Competitive Tools	Training & Development
T1	.067	.005	.851	.120
T2	.656	.112	.267	.135
T3	.252	.279	.302	.414
T4	.145	.482	.502	-.110
C1	.560	.492	-.010	.218
C2	.472	.510	.067	.273
C3	.576	.569	.032	.013
C4	.689	.372	.058	.019
D1	.542	.222	.001	.543
D2	-.041	.132	.259	.789
D3	.524	.258	-.063	.644
D4	.651	.027	.159	.537
JS1	.345	.673	.170	.019
JS2	.270	.745	.276	.150
JS3	.066	.721	.126	.384
JS4	.270	.741	.088	.237
P1	.041	.098	.867	.177
P2	.210	.339	.578	.107
P3	.694	.198	.241	.013
P4	.727	.289	-.007	.172

^a Rotation converged in 7 iterations, variance explained is 69.5%. Because of cross loadings, T4, C1, C2, C3, and D1 were deleted. T3 was deleted due to low score.

Table 2b - Descriptive Statistics, Means, Standard Deviation and Alpha

	N	Min	Max	Mean	Std. Dev	Cronbach's Alpha
Competitive Tools	99	1.00	4.00	1.9832	.76395	0.75
Practice & Operations	99	1.75	5.00	3.2071	.74833	0.79
Training & Development	99	1.00	4.67	2.8182	.75933	0.75
Job Security	99	1.75	5.00	3.6540	.76741	0.85

Table 2c: Pearson's Correlations

	Year of Service	Education	Hourly / Salary	Gender	Competitive Tools	Practice & Operations	Training & Devmt	Job Security
Year of Service	1	-.467(**)	-.51(**)	-.080	.112	.265(**)	.316(**)	.381(**)
Education	-.467(**)	1	.60(**)	-.202(*)	-.257(*)	-.131	-.288(**)	-.206(*)
Hourly vs. Salary	-.515(**)	.601(**)	1	-.127	-.260(**)	-.202(*)	-.354(**)	-.400(**)
Gender	-.080	-.202(*)	-.12	1	.072	-.034	.143	-.051
Competitive Tools	.112	-.257(*)	-.26(**)	.072	1	.342(**)	.325(**)	.409(**)
Practice & Operations	.265(**)	-.131	-.20(*)	-.034	.342(**)	1	.564(**)	.608(**)
Training & Development	.316(**)	-.288(**)	-.35(**)	.143	.325(**)	.564(**)	1	.508(**)
Job Security	.381(**)	-.206(*)	-.40(**)	-.051	.409(**)	.608(**)	.508(**)	1

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

The data was further analyzed to see whether the above results were driven by differences in gender, job category, education, or year of service. Accordingly, one-way ANOVA was performed across all 4 factors. Only statistically significant results are presented in Table 3a and Table 3b. Finally, additional analysis was conducted; GLM provides significant results for “Job Security” and “Training and Development”. These results were reported in Table 4a and Table 4b. Taken as a whole, we concluded that the Continuous Improvement program was better received by the salary respondents. Also, resistance appears to increase with year of tenure in this plant, these results were more prominent for “Job Security” than “Training and Development”.

Table 3a – ANOVA – By Year of Service

		Sum of Squares	df	Mean Square	F	Sig.
Competitive Tools	Between Groups	2.234	3	.745	1.287	.283
	Within Groups	54.960	95	.579		
Practice & Operations	Between Groups	4.027	3	1.342	2.508	.064
	Within Groups	50.853	95	.535		
Training & Development	Between Groups	5.946	3	1.982	3.724	.014 **
	Within Groups	50.560	95	.532		
Job Security	Between Groups	9.180	3	3.060	5.990	.001 ***
	Within Groups	48.533	95	.511		

*** significant at 0.001

** significant at 0.01

* significant at 0.05

Table 3b – ANOVA – by Job (Hourly vs. Salary)

		Sum of Squares	df	Mean Square	F	Sig.
Competitive Tools	Between Groups	3.879	1	3.879	7.057	.009 ***
	Within Groups	53.315	97	.550		
Practice & Operations	Between Groups	2.250	1	2.250	4.147	.044 *
	Within Groups	52.630	97	.543		
Training & Development	Between Groups	7.079	1	7.079	13.893	.000 ***
	Within Groups	49.426	97	.510		
Job Security	Between Groups	9.250	1	9.250	18.514	.000 ***
	Within Groups	48.463	97	.500		

*** significant at 0.001

** significant at 0.01

* significant at 0.05

**Table 4a: Tests of Between-Subjects Effects
Dependent Variable: Job Security**

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Corrected Model	12.876(a)	6	2.146	4.403	.001***
Intercept	80.477	1	80.477	165.128	.000***
Education	.962	3	.321	.658	.580
Hourly vs. Salary	3.714	1	3.714	7.620	.007***
Gender	.164	1	.164	.336	.563
Year of Services	2.186	1	2.186	4.485	.037*
Error	44.837	92	.487		
Total	1379.563	99			
Corrected Total	57.713	98			

R Squared = .223 (Adjusted R Squared = .172)

**Table 4b: Tests of Between-Subjects Effects
Dependent Variable: Training & Development**

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Corrected Model	9.730(a)	6	1.622	3.190	.007 ***
Intercept	45.932	1	45.932	90.341	.000 ***
Education	.356	3	.119	.234	.873
Hourly vs. Salary	1.465	1	1.465	2.881	.093#
Gender	.864	1	.864	1.698	.196
Year of Services	1.511	1	1.511	2.971	.088#
Error	46.775	92	.508		
Total	842.778	99			
Corrected Total	56.505	98			

R Squared = .172 (Adjusted R Squared = .118)

*** significant at 0.001

** significant at 0.01

* significant at 0.05

significant at 0.10

CONCLUSION

Overall, the results from this study show a bulk of the problem lies primarily with an aging and high seniority hourly workforce. Thus, one challenge for the top management is how to get these techniques effectively implemented in a workforce that has built in resistance to change because it sees using fewer workers to produce more as a negative and a long-term threat to job security? As previously stated, salaried people consistently provided higher positive ratings of continuous improvement initiatives. In addition, higher seniority was directly correlated with negative ratings. Despite the enthusiasm and the support of the CEO, results also indicate that there is a lack of committed leadership in following through the day-to-day operations at this research site.

REFERENCES

- [1] Womack, J. and Jones, D.T. (1994), "From lean production to the lean enterprise", *Harvard Business Review*, Vol. 72 No. 2, pp. 93-104.
- [2] Womack, J., Jones, D.T. and Roos, D. (1990), *The Machine That Changed the World*, Rawson Associates, New York, NY.
- [3] Steward, T. A. and O'Brien, L. (2005), "Transforming an Industrial Giant", *Harvard Business Review*, Vol. 83 No. 2, pp. 114.
- [4] Dean, J.W. and Bowen, D.E. (1994), "Management theory and total quality: improving research and practice through theory development", *Academy of Management Review*, Vol. 19 No. 3, pp. 392-418.
- [5] Brah, S.A., Wong, J.L. and Rao, B.M. (2000), "TQM and business performance in the service sector: a Singapore study", *International Journal of Operations and Production Management*, Vol. 20 No. 11, pp. 1293-1312.
- [6] Arnheiter, E. D. and Maleyeff, J. (2005), "Research and concepts: the integration of lean management and six sigma", *The TQM Magazine*, 17, pp. 5-17.
- [7] Litton, S. (2001), "Total quality management-concept articulation", available at: www.unf.edu/∼.
- [8] Antony, J., Leung, K., Knowles, G. and Gosh, S. (2002), "Critical success factors of TQM implementation in Hong Kong industries", *International Journal of Quality and Reliability Management*, Vol. 19 No. 5, pp. 551-66.
- [9] Samat, N., Ramayah, T. and Saad, N. (2006), "TQM practices, service quality, and market orientation; some empirical evidence from a developing country", *Management Research News*, Vol. 29 No. 11, pp. 713-728.

OR TURNAROUND TIME: A CASE OF HEALTHCARE PROCESS IMPROVEMENT

Michael Hall, M.D., North Shore University Hospital
Donna Moravick, RN, NP, North Shore-LIJ Health System
John F. Affisco, Hofstra University, (516) 463-5362, John.F.Affisco@hofstra.edu

ABSTRACT

This paper presents a case of the application of the Deming PDSA process improvement methodology to the problem of operating room turnaround time in the cardiac surgery operating room at a major hospital. Both the results of process analysis and process improvement are given. The results of simulation studies that investigated the impact of moving from a series to a simultaneous process indicate significant annual cost savings would be realized from this strategy.

INTRODUCTION

With the aging of the baby boomer generation and a return to an era of health care cost increases in excess of the rate of inflation, attention is again being focused on ways and techniques to improve both the effectiveness and efficiency of healthcare operations. Health care industry managers, as have manufacturing managers before them, are increasingly looking to quality management to help improve the performance of their organizations. Recognizing the need for high level understanding of quality management on the part of physicians, clinical and administrative managers, as well as for quality professionals, a major suburban teaching healthcare system, the North Shore-Long Island Jewish Health System (NSLIJ) has partnered with the Frank G. Zarb School of Business at Hofstra University to develop an MBA concentration in Quality Management. This paper reports the results of a project, developed as part of requirements of the statistical quality control course, to improve the turnaround process of the cardiac surgery operating room (OR) at one of the member hospitals of this healthcare system.

PROCESS ANALYSIS

The process selected for investigation was the process of turning a cardiac OR over for the next case. In this particular hospital there are four dedicated cardiac ORs including a dedicated staff. Seven cardiac surgeons perform about 1200 open heart procedures annually. OR time is at a premium and has a variable cost of about \$4,000 per hour after an initial fixed cost of about \$18,000. Two surgeries per day are typically scheduled in each OR, although particularly long, complex procedures may be scheduled for the whole day.

OR managers typically work hard to study surgery start time, but tend not to focus on the efficient starting of subsequent cases. This sub-optimization generally results in little, if any,

improvement in the overall efficiency and effectiveness of the utilization of these facilities. It also tends to de-motivate staff to worry about first starts since that alone has little effect on the overall progress of the day's work. Given this, OR turnaround remains a fertile area for investigation.

PROCESS DESCRIPTION

Figure 1 presents a process flow diagram for the cardiac OR turnaround process. As identified by the legend, the turnaround process starts when the first patient is removed from the OR, and is marked as the "patient out" time on the official OR records. The diagram shows the closing and transport steps prior to removing the patient from the OR. At the end of closure, the OR nurse alerts Environmental Services that the room will need to be cleaned as soon as the first patient leaves for the cardiothoracic unit (CTU) with the surgeon, anesthesiologist, physicians' assistants, and monitoring technician, with the patient attached to a portable monitor displaying blood pressure and electrocardiogram (EKG).

When the room has been cleaned, the nurses, perfusionists, monitoring technician, and anesthesiologist begin the set-up for the next case. After all is in readiness, the patient is brought into the actual OR that will be used for the procedure. This time is "patient entry" is also recorded in the official OR record form. This next patient then is placed on supplemental oxygen, and connected to the main OR monitoring system with EKG leads, arterial line placement, and other peripheral or central IV lines as needed. The anesthesiologist performs these tasks, although usually physicians' assistants and a monitoring technician will be assisting in this process. At the point where the anesthesiologist has finished his/her preparations, this is recorded as "anesthesia release time". Usually another five or ten minutes is required before the patient is prepped and draped and ready for an incision. The incision time is recorded on the official OR records when the incision is actually made by the surgeon. What is noteworthy about this process flow is that all but a few activities are carried out in series in the OR.

OPERATIONAL DEFINITIONS

The following operational definitions were agreed upon:

Turnaround Time (TT) is the time which begins when the first patient leaves the OR and ends when the next patient enters the OR. Turnaround time is made up of room cleaning time and set-up time.

- *Room cleaning time* is the time used by the Environmental Services Team to clean the room between patients.
- *Set-up time* is the time used to set-up the OR for the next patient, which includes opening instrument trays, preparing the heart-lung machine, and preparing anesthesia supplies.

Time to Incision (TI) is the time beginning when the next patient enters the room until the surgeon makes the incision. This time period is essentially "anesthesia time."

- *Release time* is part of this time, although not found to be a useful measure in this study. It is simply the time that the anesthesiologist “releases” the patient for prepping and draping, which usually takes an additional five to ten minutes.
- *Incision time* is the actual time of the surgeon’s incision, which marks the end of the period “time to incision”, and the beginning of the actual surgical procedure.

Total Turnaround Time (TTT) is the sum of the turnaround time (*TT*) and the time to incision (*TI*). The TTT represents the elapsed time that the surgeon experiences while waiting to operate upon the next patient.

In general there is a precedence relationship between most of the activities conducted during these time periods. For example, the Environmental Service workers can start cleaning while nurses break down and remove used instruments, but new instruments cannot be opened until the room cleaners are through. The turnaround time is tracked by the OR quality assurance nurse and the data are managed by an information technology staff member. The *room cleaning time* is tracked by the Environmental Services department separately using a computerized bed tracking phone system, which requires the cleaning team to call in when beginning the cleaning, and to call out when the cleaning is completed. Thus, a very accurate record is kept of the *room cleaning time*. We can safely assume that the difference between the *turnaround time* and the *room cleaning time* equals the total *set-up time*. As mentioned previously, all times except the room cleaning time are recorded on the official OR record by the room registered nurse (RN).

Difficulties experienced by the anesthesiologist in placing lines or intubating the patient will cause delays that prolong *set-up time*, since that occurs before the *incision time*, and after the *room cleaning time*. Rarely, inability to insert a Foley catheter, for instance in the case of urethral stricture, may require finding a urologist to perform an emergency flexible cystoscopy for catheter placement, which can seriously delay progress. This delay would be added to *set-up time*, again because it occurs before the incision can be made. Surgeon delays would show up as prolonged time between *release time* and *incision time*, but was not found to be a factor in this study.

DATA COLLECTION AND ANALYSIS

It was decided to measure three time periods to determine if the turnaround process is in control. Data were collected for *turnaround time*, *room cleaning time*, and *time to incision* for each of the four operating rooms for the months of May, June, and July, 2006. *Total turnaround time* could then be calculated by adding the *turnaround time* and *time to incision* for each case. A total of 79 cases were found to qualify as second open heart cardiac surgery cases out of a total of 361 cardiac surgery service cases in that three month period. In addition, since Environmental Services had a separate time tracking system, data for nine months, including May, June, and July of 2006, was available for analysis.

Since all data represented individual observations of the measurable variable time, Individual and Moving Range (I-MR) variables control charts were chosen to monitor the performance of

the turnaround process [2]. Minitab 14 was used to construct and assist in the analysis of the control charts and for process capability.

Turnaround Time (TT)

An initial look at the turnaround data for the three month period showed an anomaly. Results of discussions with clinical managers indicated that the data were adjusted and all turnaround times in excess of 75 minutes were considered outliers and simply eliminated. This was essentially a standard that existed across the health care system. With the cooperation of senior management, the investigators were able to obtain the complete set of raw data.

Turnaround was calculated for 79 cases over the three month period. Any emergencies or add-ons were excluded from the analysis to allow for a better sense of process capability. Figure 2 presents the I-MR charts for turnaround time. The moving range chart shows points 57, 58, 64, and 65 more than 3σ from the centerline. In addition, points 66 through 78 are on the same side of the centerline. This process is clearly not in control. Although the individuals chart is relatively meaningless in face of an out of control moving range chart, one must note that when outliers are included, the average turnaround time is 72.2 minutes with an upper control limit of 142.5 minutes (3σ).

To determine if excluding the health care policy outliers of turnaround times in excess 75 minutes would bring the process into control, a second set of I-MR charts was constructed. Figure 3 presents these charts. This left 57 out of 79 cases. For this three month period, 22 cases or 27.85% of the total three month turnaround times were "outliers". This I-MR chart also demonstrated data out of control, even with the "outliers" excluded. These results led the investigators to conclude that actual turnaround time (TT) and turnaround time with "outliers" excluded are both out of control. Average TT is 72.2 minutes compared to 60.12 minutes with "outliers" excluded. Given average TT = 72.2 minutes, excluding all data over 75 minutes seems unwise if a true picture of process is desired.

Environmental Services – Room Cleaning Time

Figures 4, 5, and 6 present the control charts for cleaning time for May, June, and July 2006 respectively for all four ORs. Test Results for the May 2006 I Chart for ES shows 4 out of 5 points more than one standard deviation from center line (on one side of CL). The mean cleaning time for May was 16.85 minutes with an upper control limit of 28.44 minutes. However, these data are in better control than the overall turnaround data for May 2006. It appears that only two points barely go into the 2σ or B zone, and none beyond the 3σ upper limits of control. It should be noted that data for late night and weekend terminal cleaning and small spills were eliminated from the data. The I-MR Chart for June shows a process in control with a mean of 17.81 minutes and upper control limit of 32.57 minutes. The times are similar to May. The MR Chart for July 2006 is clearly in control. It is interesting that on the I chart for July, Minitab identifies some points which violate its test 5 and 6. However, both tests involve multiple points on the low side of the centerline, which is very desirable. The centerline is at an average of 16.12 minutes with an upper control limit of 24.54 minutes. All late night and weekend terminal cleaning times, as well as small spill cleanup times, were removed from the data as before.

Room cleaning data analysis shows environmental service performance is in control. During the busiest OR month (June) Environmental Services provided fast cleaning times. The Environmental Services department was adequately staffed and motivated to meet all challenges during the study period. Therefore, the problem with turnaround time must lie in the set-up time component.

Time to Incision

Time to Incision (TI) data shows a process component out of control. Mean Time to Incision is 66.4 minutes with a 1σ limit of 85.4 minutes, 2σ limit of 104.4 minutes, and a 3σ limit of 123.5 minutes. Average time from when first patient A leaves to the time of second patient B incision is $72.2 (TT) + 66.4 = 138.6$ minutes. If only 1σ prevails, the surgeon waits $95.7 (TT) + 85.4 = 181.1$ minutes (about three hours) to make the next incision!

PROCESS IMPROVEMENT

As a way to validate the findings of the initial process analysis, benchmarking data for OR turnaround time was sought. A benchmarking study of 22 procedures in several different hospitals including 7,664 cases, 525 of which were coronary artery bypass graft surgery, over the four-year period 1996-1999 published in OR Manager Magazine [3] indicates that the average OR set-up time, in an institution performing large volumes of cases, is between 35 and 45 minutes, and the average cleaning time is between 15 and 20 minutes. While the cleaning times are remarkably similar to the data from the team's process analysis, the benchmark set-up times are about 50 to 75 percent of the average times discovered in the institution of our study. This served to further confirm that the institution was experiencing a serious problem in OR turnaround time as identified in the process analysis which can be traced to set-up time.

As indicated in the process flow, essentially all activities are carried out in series in the operating room. This clearly contributes to the excessive turnaround time in the ORs. Other institutions have designed processes that conduct some of the set-up activities in parallel outside of the OR. With this knowledge in hand, the investigators proceeded to develop a process improvement program. Deming's Plan-Do-Study-Act (PDSA) methodology was used as the framework for this improvement cycle [4]. Based on information gleaned from benchmarking sources, mean times with upper and lower acceptable limits were defined to allow determination of process capability [5], [6]. Process capability graphs were constructed using actual data from cleaning time (Fig. 8), turnaround time (Fig. 9), time to incision (Fig. 10), and total turnaround time (Fig. 11). Clearly, only the cleaning time was in control, including the busiest month of the period studied.

Plan

The plan phase began with a presentation of the results of the process analysis to the Deputy Executive Director of the Institution. The result of this meeting was a go ahead to develop a process improvement program. A follow up meeting was held with the multidisciplinary team that is responsible for OR turnaround. At this meeting, the concept of simultaneous versus

parallel processing was introduced. Discussion centered on the logistics of placing monitoring lines and other appropriate pre-induction equipment in or on the patient in a dedicated area (other than the cardiac OR) during the cleaning and set-up of the cardiac OR for the next case. The constraints were the physical space to allow insertion of these monitoring lines. As a contingency plan, other areas were identified as potential space for line insertion by anesthesia. The recovery room (RR) was discussed as a possible location. However, it was determined that RR reaches its full capacity early in the day. A second suggestion was made to provide space in the cardiothoracic unit (CTU), which is not congruent to the OR. At this time the nursing director of the OR informed the group that there are plans for an area that would allow same day admission patients to have central line insertions. A suggestion was made to allocate space to allow insertion of central lines, arterial lines, and so forth, for the more complex surgery patients. Of course, it appears that the optimal approach would be the construction and allocation of pre-induction rooms as a holding area dedicated to anesthesia, which was done at the Cleveland Clinic many years ago. Barring this possibility, a larger pre-induction area adjacent to the OR would be the next best approach.

Do

This phase requires that planned experiments be conducted to determine how best to implement the plan established in the previous stage. Since construction and testing of new facilities is impractical in the short term, it was decided to conduct a simulation experiment designed to compare the performance of the existing series process with the proposed simultaneous (parallel) process. A computer model of each process was built using EXTEND 6 (Figs. 12 & 13). The result of the simulation indicated significant decreases in both mean turnaround and standard deviation turnaround time for the new parallel process, when compared to the old serial process. Specifically, mean total turnaround time (TTT) decreased from 138.99 to 79.76 minutes, a reduction of 42.6% while standard deviation decreased from 16.91 to 8.477 minutes or 49.9%. When the simultaneous data were plotted on I-MR charts, the process was shown to be in statistical control.

Study

The analysis completed in the Do phase clearly supports the hypothesis that simultaneous processing would go a long way toward improving the performance of the cardiac OR turnaround process. At the present time simultaneous processing is constrained by the availability of space required to conduct certain set up activities, notably anesthesia, in parallel. Further, new space would require new construction which is clearly a long term proposition. Realizing this, attention was turned to what could be done in the short-term to improve OR turnaround. Attention was then turned to a short and intermediate term plans.

This short term plan includes the following:

- Procedures for improved communications between environmental and nursing staffs.
- Investigate possibility of safe overlap of cleaning and set up.
- Observation of present practices by unbiased staff observes to improve the quality of data that are collected.

- Identify “nursing leaders’ assigned to each of the four cardiac ORs and involve them in the improvement program.
- Track OR turnaround time per room. Post these times so that they may be viewed by all staff. Be sure “outliers” are counted!
- Consider implementing an incentive program. The cardiac OR with the most improved turnaround time will receive some award.
- Look at staff scheduling and if necessary, hire one or two OR technicians to allow initiation of set up in anticipation of the next surgical procedure. Schedule staff shifts so that a technician will be available during peak surgical times to float to the room that needs to be turned around for the next case.

In the intermediate term:

- Improve logistics with redesigned case carts. Create separate opening trays, procedural trays, and closing trays based on surgeon’s specifications.
- Develop a single tray with standardized equipment for all surgeons
- Consider the purchase of a tracking device for actual turnover and time to incision.

Act

In this phase the plan should be integrated into the process. Clearly the simultaneous process can best be affected by investing in the design and construction of one or more rooms dedicated mainly to pre-induction and/or administering of anesthesia to patients. Estimates indicate that a minimum savings of \$1.28 million dollars can be realized annually if a parallel process is implemented in the cardiac OR. If the entire OR can benefit from parallel processing, more savings will accrue, more OR utilization and throughput will be realized, and this will improve employee retention, patient and surgeon satisfaction, and hospital contribution margin.

As is the case in any successful organization, the deployment of resources for this purpose must be evaluated against other potential uses in the hospital. While a decision on this investment is considered, the short and intermediate term plan is in the process of being implemented.

CONCLUSION

This paper presents the results of a unique process improvement project aimed at improving the performance of cardiac operating room total turnaround time at a major tertiary care hospital. The major conclusion is that both long and highly variable set-up times and anesthesia times, result in unacceptably long total turnaround times. A root cause of this is the series process that is presently used for cardiac OR procedures. The major recommendation of this study is to redesign the process to allow certain activities to be conducted in parallel, thus reducing the time required in the OR for the patient.

The complete version of the paper including references and figures is available from John F. Affisco.

THE MEDICARE DRUG BENEFIT: BUDGET BUSTER OR BENEFICIARY BOON?

Laura Forker, University of Massachusetts Dartmouth, lforker@umassd.edu, DIS Dep't.,
University of Massachusetts Dartmouth, 285 Old Westport Rd, N. Dartmouth, MA 02747-2300

ABSTRACT

When the Medicare Part D prescription drug legislation was being developed, lawmakers decided to specifically prohibit the Medicare program from bargaining with pharmaceutical companies to secure lower drug prices. This controversial decision took the responsibility for moderating drug prices away from the Medicare program and, instead, placed it in the hands of private drug plans. Two years after the implementation of Part D, the consequence of this decision is that private plans have failed to deliver low prices. The Medicare drug benefit increasingly appears to be a budget buster. Escalating drug prices appear to be unsustainable over the long term.

Keywords: Health care, Medicare, policy implications

INTRODUCTION

The Medicare Prescription Drug, Improvement, and Modernization Act of 2003 (MMA) created a new market for pharmaceutical companies to offer prescription drugs. Part D of the legislation uses private, stand-alone drug plans that compete in 39 regions (34 covering the States and five others, each covering a territory) to make prescription drug benefits available to beneficiaries covered under traditional Medicare [10]. As of January 2007, 43 million beneficiaries obtained their prescription drug coverage through Medicare [15]. Of this total, 23.9 million people (56%) were enrolled in Medicare Part D plans [15]. Of those beneficiaries in Part D plans, 11 million (26%) utilized stand-alone prescription drug plans (PDPs); 6.3 million (15%) were dual eligibles (individuals who qualify for both Medicare and Medicaid) in PDPs; 6.7 million (16%) were Medicare Part C participants enrolled in Medicare Advantage drug plans (MA-PDs), including 0.5 million dual eligibles; 10.3 million (24%) had creditable coverage through employer/union plans (including FEHB and TRICARE coverage); 4.9 million (11%) had creditable coverage from the Veterans Administration, Indian Health Service, state pharmaceutical assistance programs, or employer plans for active workers; and 4 million (9%) were estimated to have no creditable coverage [15]. Unlike most forms of insurance, the Medicare Part D prescription drug program has a coverage gap in the middle known as the “doughnut hole”. After beneficiaries reach their initial limit of total drug expenses, they have no prescription drug coverage until their total drug expenses reach a catastrophic threshold for the year. While beneficiaries are in the doughnut hole, they must continue to pay monthly plan premiums, although they do not receive any drug benefits. These monthly premiums have been rising steadily and are predicted to rise 21% on average in 2008 [22]. Only after beneficiaries have spent thousands of dollars of their own money to get out of the doughnut hole, in addition to their monthly premiums, does their coverage resume. Proponents of Part D have suggested that beneficiaries seeking full coverage

through the doughnut hole should switch from traditional Medicare to a Medicare Advantage managed care plan that provides such coverage. While providing prescription drug coverage, most Medicare Advantage plans impose limits on which doctors their members can see or which hospitals they can use. Beneficiaries are therefore at risk of losing access to their chosen health care providers. They may also face higher cost sharing for some medical services than they would have under traditional Medicare [16].

THE MEDICARE DRUG BENEFIT: BUDGET BUSTER?

When the Medicare Part D prescription drug legislation was being developed, lawmakers decided to specifically prohibit the Medicare program from bargaining with pharmaceutical companies to secure lower drug prices. This controversial decision took the responsibility for moderating drug prices away from the Medicare program and, instead, placed it in the hands of private drug plans. Two years after the implementation of Part D, the unfortunate consequence of this decision is that private plans have failed to deliver low prices. The Medicare drug benefit increasingly appears to be a budget buster. Drug prices set by private Part D plans significantly affect premiums and how much beneficiaries end up paying out of pocket overall. These drug prices also have a direct effect on taxpayers who shoulder approximately three-fourths of the costs of the Part D program [18]. Specifically, the rule that prohibits Medicare from establishing its own plan, and negotiating price discounts directly with the pharmaceutical industry, plus the added administrative costs associated with supplying drug assistance, have added hundreds of billions of dollars to the cost of the Medicare program [1]. Other costs that are more difficult to document but are substantial nevertheless include the time that beneficiaries, health care providers, and pharmacists spend analyzing various policy options due to the excessive complexity of alternative plans [1]. Part D plans can impose quantity limits, require enrollees to pay large co-payments, and restrict access through utilization review or prior authorization [6]. Prescription drugs for dual eligibles are also more expensive through Medicare Part D than they were previously under Medicaid [17].

THE MEDICARE DRUG BENEFIT: BENEFICIARY BOON?

As Congress debated the MMA in 2003, President Bush stated, “all seniors should be able to choose the health care plan that best fits their needs without being forced into an HMO” [16]. This is why stand-alone Part D drug plans were established - to supplement traditional Medicare rather than to limit drug coverage to Medicare Advantage managed care plans. Pharmaceutical manufacturers, wholesale distributors, pharmacies, and insurance companies have all grown profitable as a result of the added number of prescription drug subscribers and also because of their increased ability to charge higher prices for prescription drugs to a wider population.

Pharmaceutical Manufacturers

Pharmaceutical manufacturing is a large global industry. In 2003, worldwide pharmaceutical industry sales totaled \$491.8 billion, an increase in sales volume of 9 percent over the preceding year [11]. The United States represents the largest single national market for pharmaceuticals, accounting for 44 percent of global industry sales in 2003 (\$216.4 billion), an increase of

approximately 12 percent from the previous year's figure [12]. After a decade of significant mergers and acquisitions by drug companies, the ten largest pharmaceutical manufacturers accounted for almost 60 percent of total U.S. sales in 2004 [14]. In the six months after the Part D drug legislation was enacted (January 1 - June 30, 2006), the profits of these drug companies ballooned by more than \$8 billion [20]. The Medicare Part D program was partly responsible for this swell in industry profits because of the opportunities cited above. The inability of the government to negotiate price discounts directly with the pharmaceutical industry resulted in higher drug prices for recipients. The transfer of dual eligibles from Medicaid to Medicare Part D led to higher per capita drug expenses for the poor. And the availability of a drug benefit where none had existed before culminated in an increase in drug utilization by all seniors and disabled. The transfer of the 6.1 dual eligibles (those qualifying for both Medicare and Medicaid) has resulted in an especially sharp price increase for pharmaceuticals for this population due to the change in service delivery from Medicaid, where drug manufacturers were required to provide the federal government with the lesser of either their lowest price available or a 15.1% markdown off the "average manufacturers price" of the drug [20]. When the Part D legislation took effect, the 6.1 dual eligibles were diverted to the Medicare program and the drug manufacturers were no longer required to sell their drugs at their lowest prices; the resulting price increases have been especially burdensome for those who can least afford them [20]. Congressional estimates have projected that price escalation for pharmaceuticals could lead to industry profits of more than \$2 billion in 2006 and over \$30 billion by 2016 [20].

Consumers and Taxpayers

Consumers and taxpayers are paying for a substantial portion of the Part D program since Medicare uses tax revenues to pay for roughly 75 percent of Part D [7]. Higher prices also affect the elderly and people with disabilities covered under Part D, who must pay a considerable portion of their drug costs out of pocket. For example, in 2008 Humana is planning to hike the average premium for its basic plan nationwide by 69%; UnitedHealth Group is expanding its AARP Savers plan premium by 87% [27]. Part D enrollees must also continue paying monthly premiums while in the doughnut hole while receiving no prescription drug coverage [7]. The MMA legislation prevents Medicare from using its considerable leverage, forbidding it from negotiating to secure lower drug prices while simultaneously overpaying private plans that fail to save money. "The result is windfall profits for drug and insurance companies, and lighter wallets for American taxpayers, seniors, and people with disabilities" [7]. The Congressional Budget Office (CBO) estimates that the program will cost \$850 billion in its first 10 years of operation from 2006 to 2015 [4].

WHAT CAN THE FEDERAL GOVERNMENT LEARN FROM STATE EXPERIMENTS AND OTHER COUNTRIES' SYSTEMS?

The affordability of prescription drugs for older and disabled persons has emerged as an urgent and high profile state health policy issue. There are over 30 states at this time that have instituted programs that provide some form of assistance to older and disabled residents who lack drug coverage. These programs differ substantially from state to state, as some offer direct benefits, some are targeted to lower-income beneficiaries, while others reach up to more moderate-income

levels. Some provide unlimited benefits, while others impose annual benefit caps on enrollees or restrict coverage to certain classes of drugs [5]. State-sponsored pharmacy assistance programs (SPAPs) provide prescription drug coverage for low-income, older, and disabled persons who are not eligible for Medicaid and who may have no other drug coverage [9]. Faced with increasing costs and mounting budget pressures, SPAPs have experimented with a variety of cost-control measures such as [9]:

- Substitution of generic medications for brand-name products.
- Prior authorization (i.e., the state reviews a prescription before it is dispensed).
- Seeking improved manufacturer rebates, differential co-payments for preferred and nonpreferred medications.
- Restricted formularies.
- Use of pharmacy benefit managers and administrators.
- Reimbursement of the state by other insurance plans

Other state-based initiatives might include tax credits for drug purchases, buyers' clubs, and purchasing cooperatives [2].

Most countries other than the United States regulate the prices of drugs, whether directly through price controls, indirectly through limits on reimbursements under social insurance programs, or indirectly through profit controls imposed on drug companies [6]. Prices in other nations generally reflect the lower incomes in some countries and the politicized nature of most foreign health care systems that rely on price controls and access restrictions [2].

POLICY RECOMMENDATIONS TO MAKE DRUGS MORE AFFORDABLE

The Medicare drug benefit is likely to undergo administrative and potentially legislative changes in its early years, as evidence of its costs and benefits emerge. Several approaches to controlling program costs could include: changing eligibility criteria, modifying the scope of benefits, modifying the level of consumer cost-sharing, leveraging market share to negotiate or mandate pharmacy discounts and manufacturer rebates, and tightly administering the program [5]. More potent methods could focus on indirect price controls, coerced discounts, dilution of patent rights, and re-importation of U.S.-manufactured drugs from lower-priced foreign markets to force down prices [2].

Some manageable policy recommendations to lower the price of prescription drugs include limiting direct-to-consumer (DTC) advertising, improving the patient-physician relationship through greater information sharing by physicians to patients, and creating a program similar to the Veterans Administration (VA) system.

Proponents of DTC advertising argue that it informs consumers about important, treatable health conditions and encourages doctor patient communication. Critics say that this type of advertising contributes to rising drug costs and leads patients to demand unnecessary or inappropriate medications [13]. Drugs that are heavily advertised to consumers typically rank high in sales: 6 of the top 10 drugs advertised through DTC promotions were among the top 20 drugs in dollar

sales and in the number of prescription dispensed in 2000 [13]. DTC advertising produces a significant return for the pharmaceutical industry as well, where every additional \$1 the industry spent on DTC advertising in 2000 yielded an additional \$4.20 in sales [13]. This form of advertising is banned in several European countries and could be curtailed in the United States as well. DTC advertising has transformed many once-passive patients into inquiring and demanding consumers and can affect physician choice of specific medications prescribed and modes of delivery used [14]. Physicians need to discuss DTC advertising with their patients, and they should be prepared to explain why a particular advertised drug is not appropriate, if needed [2].

The patient-physician relationship itself is also in need of improvement, as 40 percent of seniors responding to a survey reported not adhering to their doctor's orders regarding their medication regimens [21]. Patients who discussed their medications with their physicians, the researchers found, were more likely to be switched to lower-cost drugs [21]. This communication gap needs to be addressed, as knowledge is particularly important in what has traditionally been a physician-led process: patients can ask better questions and make better judgments if they know some of the options available [2]. Patients and physicians together should weigh the pros and cons of brand name versus generic drugs to determine the best treatment, since generic drugs are typically less expensive than brand-name drugs and are often suitable substitutes.

The legislation that created Part D strictly prohibits Medicare from playing any part in negotiating drug prices with pharmaceutical companies. Instead, it relies solely on the power of the private market to secure lower prices. This stands in stark contrast with other government agencies, such as the VA system, which uses the leverage it derives from the millions of veterans it serves, to negotiate lower drug prices [7]. Part D plans can require large co-payments, impose quantity limits, and restrict access through prior-authorization requirements and utilization review due to inadequate leverage on the part of consumers and institutional buyers. In contrast, the VA charges only a small co-payment per prescription and rarely imposes restrictions on the use of drugs [18].

CONCLUSION

Pharmaceutical products provide medical benefits to patients as well as economic benefits to society by reducing other medical expenditures, lowering the rate of hospitalizations, minimizing surgeries and other invasive medical treatments, and reducing the mortality rates from many diseases [3]. This research illuminates the difficulties facing many elderly Americans who are using prescription drugs more than ever before: 32 percent of seniors report three or more chronic conditions, while 41 percent use five or more prescriptions [27]. Of those who use multiple medications, 54 percent have two or more prescribing physicians and 36 percent use two or more pharmacies [27]. Thirty-two percent spend more than \$100 per month out-of-pocket on medications [27].

The only way to cultivate new drugs is to invest heavily in research and development. The \$30 billion spent annually by U.S. drug makers tends to be high, as they create a "virtuous cycle" that encourages more R&D to create newer medicines [3]. A study of drug development calculated that "between the time research begins to develop a new prescription medicine, until it receives

approval from the FDA to market the drug in the United States, a drug company typically spends \$802 million over the course of 10 to 15 years.” [25].

The future is likely to see even greater medical advances, as the biotech revolution offers particular promise, with the prospect of tailoring drugs to patients’ individual genetic characteristics. People continue to use prescription drugs, even while complaining about the cost, because of the benefits they receive when drugs are used properly. For consumers who need these new drugs, “Part D plans without meaningful coverage through the doughnut hole is woefully incomplete” [22]. Nearly two years of implementing Part D have shown that private plans have failed to deliver lower prices. Drug prices set by private Part D plans have expanded most premiums (some significantly) and driven up the amount beneficiaries must pay out of pocket, disappointing the seniors, taxpayers, and the 43 million beneficiaries who want lower prices. These vulnerable beneficiaries need meaningful coverage through the doughnut hole, but such coverage is becoming either increasingly expensive or unavailable. Legislation must be changed so that the government can use its considerable bargaining power to lower Part D drug prices and to allow competition from *all* sources, including re-importation of American-made drugs, if rising drug prices (and overall costs to Part D beneficiaries) are to be curtailed.

References available upon request from Laura Forker.

ON THE INSPECTION ERRORS IN THE DICHOTOMOUS INSPECTION PROCEDURE

Young H. Chun

E. J. Ourso College of Business, Louisiana State University, Baton Rouge, LA
(225) 578-2506; chun@lsu.edu

ABSTRACT

Suppose that a target is represented by a point on a given interval. When we test a point in the interval, we may obtain directional information about the location of the target. Looking up a telephone directory or an English dictionary is a good example of the so-called “dichotomous search” in everyday life. The dichotomous search has been widely applied to the problem of inspecting a sequence of items when the process failure rate is constant. From the Bayesian point of view, we propose an inspection model with inspection errors.

Keywords: Quality control, Inspection errors, Search theory, Bayesian analysis

1. INTRODUCTION

Consider a production process that produces a batch of n items or units. At the beginning of a production run, the production process always starts “in control” and produces only “non-defective” items. If the process goes “out of control”, it produces only “defective” items. Thus, the entire sequence of n items can be divided into two disjoint sets: a stream of non-defective items, followed by a stream of defective items.

When we inspect one of the n items in the sequence, we can guess the exact timing of the process shift. If the tested item is defective, for example, the process went out of control before the defective item was produced. In such a case, we must reject all the items produced after the tested item. If the tested item is non-defective, on the other hand, all the items produced before it in time should be accepted without further inspections. With repeated inspections, we can detect more defective items and identify more non-defective items in the sequence.

Many authors have dealt with the problem of classifying all the items in the sequence with a minimum number of inspections. For the production process with a constant failure rate, for example, Hassin [2] proposed the optimal procedure that minimizes the total number of inspections. Recently, attention has been shifted to the economic design of a sequential inspection procedure that minimizes the total cost. Based on several cost factors associated with inspection and classification activities, Raz, Herer, and Grosfeld-Nir [6] developed an inspection procedure in which we are permitted to terminate the inspection process earlier if it is more cost-effective.

Recently, Sheu, Chen, Wang, and Shin [7] simply added inspection errors to the cost-minimization model. Wang [8] pointed out major flaws in [7] and proposed another inspection, which we believe still contains critical errors. Those errors can be easily explained from the Bayesian point of view.

2. BAYESIAN DICHOTOMOUS SEARCH

Consider a sequence of n items that have been produced from the same production line and arranged in chronological order. Each item is either “defective” or “non-defective”. We define a binary variable Y_i to represent the quality of the i th item in the sequence:

$$Y_i = \begin{cases} 1 & \text{if the } i\text{th item is non - defective} \\ 0 & \text{if the } i\text{th item is defective.} \end{cases} \quad (1)$$

Once the production process shifts into the “out of control” state, it produces only defective items. Thus, we assume that $Y_j = 0$ if $Y_i = 0$ for all $i < j$. Likewise, $Y_i = 1$ if $Y_j = 1$ where $i < j$. Consequently, the set of binary variables Y_i is divided into two disjoint sets: a set of ones followed by a set of zeros. From the search theoretic point of view, let us define the “target” as the first defective item in the sequence.

Finding the target with directional information has been known as a “dichotomous” search in operations research. Once we test a trial point within the interval, we know that the target is located on the right-side or left-side of the trial point, and we can narrow down the interval on which the target lies. The bisection method for finding the maximum or minimum point in a single-variable optimization problem is a good example of the dichotomous search. Looking up a telephone directory or an English dictionary is another example of the dichotomous search in everyday life.

In some cases of dichotomous search, we may have a prior knowledge about the location of the target, and the prior knowledge is represented as a probability distribution in Bayesian analysis. Let T_i denote the event that the i th item in the sequence is the target or, in other words, the stream of defective items begins with the i th item. Let $P[T_i] = p_i$ be the “prior” probability that the i th item is the target, where $i = 1, 2, \dots, n$. For notational convenience, let $\pi_j = p_1 + p_2 + \dots + p_j$ be the *cumulative* probability that the target is one of the first j items in the sequence. If all the items in the sequence are non-defective, the cumulative probability π_n is less than one. Let $p_{n+1} = 1 - \pi_n$ be the probability that there is no target in the sequence of n items; i.e., all the items in the sequence are non-defective.

In early years, the discrete *uniform* distribution was the most popular choice as the *a priori* distribution of the target [6]; chances are equally likely that the target is hidden in one of the n locations. In such a case, it can be shown that the bisection method is the optimal search strategy that minimizes the expected number of inspections. Hassin [2] is the first who proposed a *geometric* prior, assuming that the production process has a constant failure rate. His geometric prior was later extended by He, Gerchak, and Grosfeld-Nir [3], who considered the possibility that the target may not exist in the sequence. Raz, Herer, and Grosfeld-Nir [6] also used the geometric prior for their cost-minimization inspection model. But the geometric prior distributions considered in [2] and [3] are slightly different from each other.

3. TWO TYPES OF GEOMETRIC PRIOR

There are two types of inspection plans in dichotomous search. In the “on-line” inspection plan, on one hand, every n th item is inspected while the production process is still running. The production process stops when we detect a defective item. In such a case, the last item in the sequence of n items is

obviously defective and we are interested in the search procedure that identifies more defective items in the sequence of n items.

In the “off-line” inspection plan, on the other hand, the production process has already produced a batch of n items. We assume that the on-line inspection is not feasible because, for example, the inspection time of an item is longer than its production time. In such a case, the problem of interest is how to detect, if any, defective items in the sequence with a minimum number of inspections. Thus, there are two different geometric prior distributions for the production process with a constant failure rate.

First, for the *off-line* inspection model, He, Gerchak, and Grosfeld-Nir [3] used the geometric distribution as follows:

$$p_i = \begin{cases} \phi^{i-1}(1-\phi) & \text{for } i = 1, 2, \dots, n \\ \phi^n & \text{for } i = n + 1 \end{cases} \quad (2)$$

where ϕ is a parameter, which is interpreted as the probability that the process remains in-control after producing each item. From (2), the cumulative distribution function is

$$\pi_i = \sum_{k=1}^i p_k = 1 - \phi^i, \text{ for } j = 1, 2, \dots, n. \quad (3)$$

There is a possibility that there is no defective item in the sequence; this probability is $p_{n+1} = \phi^n$.

Second, the geometric prior considered in Hassin [2] is for the *on-line* inspection plan. Because at least one of the items is defective, the geometric prior is truncated after the n th item as follows:

$$p_i = \begin{cases} \phi^{i-1} \frac{1-\phi}{1-\phi^n} & \text{for } i = 1, 2, \dots, n \\ 0 & \text{for } i = n + 1 \end{cases} \quad (4)$$

and its cumulative distribution function is

$$\pi_i = \sum_{k=1}^i p_k = \frac{1-\phi^i}{1-\phi^n}, \text{ for } j = 1, 2, \dots, n. \quad (5)$$

Raz, Herer, and Grosfeld-Nir [6] considered both types of geometric prior in the economic design of optimal inspection procedures. The popularities of the geometric and truncated geometric distributions in dichotomous search are due to their unique properties that significantly simplify the dichotomous search procedures. Those properties will be reviewed in the next section.

4. GEOMETRIC PRIOR WITH INSPECTION ERRORS

In the presence of inspection errors, we need to define another indicator variable X_i as follows:

$$X_i = \begin{cases} 1 & \text{if the } i\text{th item is tested non - defective} \\ 0 & \text{if the } i\text{th item is tested defective.} \end{cases} \quad (6)$$

With the indicator variables X and Y , the misclassification errors α and β in Sheu, Chen, Wang, and Shin [7] are defined as shown in Table 1.

Table 1. Misclassification errors in the dichotomous inspection procedure

		Inspection Result	
		X=1 (Non-defective item)	X=0 (Defective item)
Fact	Y=1 (Non-defective item)	1 - α	α
	Y=0 (Defective item)	β	1 - β

Suppose that we decide to continue the search process and inspect the k th item in the sequence. Its inspection result X_k is either non-defective or defective. Depending on the inspection result, the posterior probabilities can be obtained as follows.

Table 2. Posterior probabilities for the geometric priors in the presence of inspection errors

Prior distribution, p_i		Geometric $\phi^{i-1}(1-\phi)$	Truncated geometric $\phi^{i-1}(1-\phi)/(1-\phi^n)$
If the k th item is non-defective ($X_k=1$)	P[$X_k=1$]		$\beta(1-\phi^k) + (1-\alpha)\phi^k$
	i	1, 2, ..., k	$\frac{\beta\phi^{i-1}(1-\phi)}{\beta(1-\phi^k) + (1-\alpha)\phi^k}$
		$k+1, k+2, \dots, n$	$\frac{(1-\alpha)\phi^{i-1}(1-\phi)}{\beta(1-\phi^k) + (1-\alpha)\phi^k}$
		$n+1$	$\frac{(1-\alpha)\phi^n}{\beta(1-\phi^k) + (1-\alpha)\phi^k}$
If the k th item is defective ($X_k=0$)	P[$X_k=0$]		$(1-\beta)(1-\phi^k) + \alpha\phi^k$
	i	1, 2, ..., k	$\frac{(1-\beta)\phi^{i-1}(1-\phi)}{(1-\beta)(1-\phi^k) + \alpha\phi^k}$
		$k+1, k+2, \dots, n$	$\frac{\alpha\phi^{i-1}(1-\phi)}{(1-\beta)(1-\phi^k) + \alpha\phi^k}$
		$n+1$	$\frac{\alpha\phi^n}{(1-\beta)(1-\phi^k) + \alpha\phi^k}$

First, the probability that the k th item is tested non-defective can be shown to be

$$P[X_k = 1] = \beta \sum_{i=1}^k p_i + (1 - \alpha) \sum_{i=k+1}^n p_i + (1 - \alpha)p_{n+1} = \beta\pi_k + (1 - \alpha)(1 - \pi_k). \quad (7)$$

In such a case, the posterior probabilities of the target are summarized in the upper half of Table 2 for the two types of geometric prior in (2) and (4).

Second, the probability that the k th item will be tested defective is

$$P[X_k = 0] = (1 - \beta) \sum_{i=1}^k p_i + \alpha \sum_{i=k+1}^n p_i + \alpha p_{n+1} = (1 - \beta)\pi_k + \alpha(1 - \pi_k). \quad (8)$$

When the k th item is tested defective, the posterior probabilities are also displayed in the lower half of Table 2 for the two types of geometric prior in (2) and (4).

Consider the sequence of $n=10$ items. Suppose that the prior distribution of the target is a geometric distribution with $\phi = 0.8$, and the inspection errors in Table 1 are $\alpha = 0.1$ and $\beta = 0.2$. When we inspect the fourth item in the sequence, the resulting posterior distributions are displayed in Figure 1 corresponding to the inspection result X_4 .

As shown in Figure 1, (i) the interval of uncertainty does not narrow down regardless of the inspection result, (ii) the posterior distribution is no longer the geometric distribution, and (iii) the state of inspection process should be represented not by the interval of uncertainty, but by the vector of posterior probabilities.

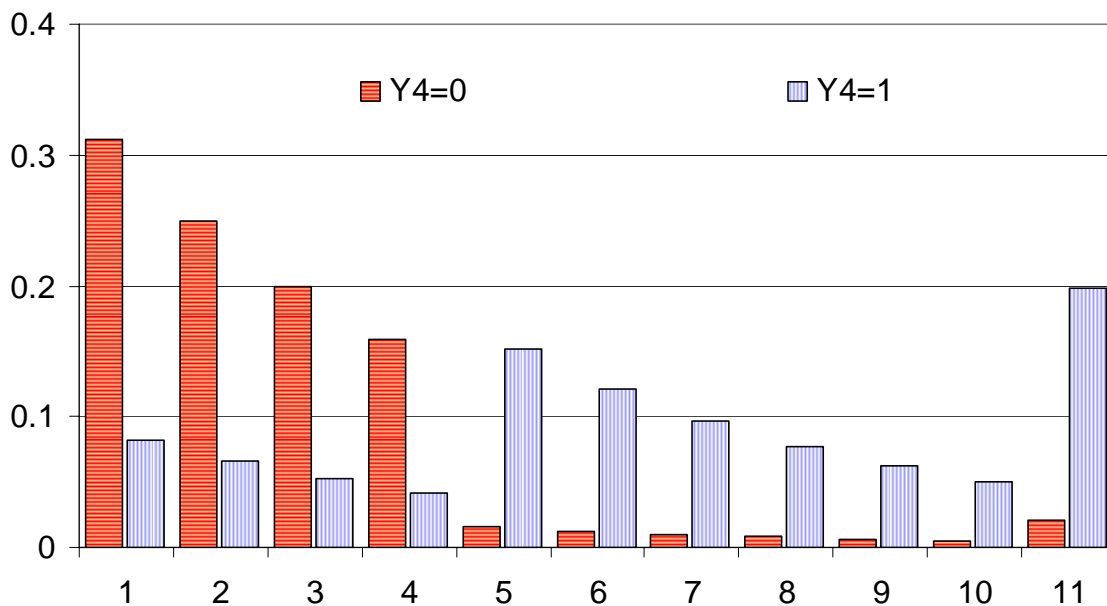


Figure 1. Posterior distributions of the target corresponding to the inspection result of the fourth item ($\alpha = 0.1$ and $\beta = 0.2$).

5. CONCLUDING REMARKS

In Table 2, we derived the correct form of posterior distributions, with which we can easily formulate a dichotomous inspection model that minimizes the expected total cost. Alternatively, we may derive an entropy-based inspection model as in [6]. As shown in Figure 1, the location of the target becomes more certain as we continue the inspection process, and the amount of certainty in the probability distribution can be measured by the “information entropy”. Thus, based on the posterior probabilities in Table 2, we can calculate the entropy for a given state of inspection process.

The dichotomous search has been applied to a wide variety of managerial decision problems, such as locating a flaw in a underground cable or pipe, constructing a binary search tree for use in computer file search, arranging items in an alphabetical order, locating nerve endings, identifying the defective items in a mass-production process, finding a proper re-crystallization temperature in physics, locating the end of a tree root, and so on (Berry and Mensch [1]). The posterior distributions derived in the paper would help develop an optimal search strategy for such a dichotomous search in the presence of inspection errors.

REFERENCES

- [1] Berry, D. A., Mensch, R. F. “Discrete search with directional information”. *Operations Research* 34, 1986, 470-477.
- [2] Hassin, R. “A dichotomous search for a geometric random variable”. *Operations Research* 32, 1984, 423–439.
- [3] He, Q-M, Gerchak, Y. and Grossfeld-Nir, A. “Optimal inspection order when process failure rate is constant”. *Int J Reliability Qual Safety Eng* 3, 1996, 25–41.
- [4] Herer, Y. T., and Raz, T. “Optimal parallel inspection for finding the first nonconforming unit in a batch – An information theoretic approach”. *Management Science* 46, 2000, 845-857.
- [5] Morris, M. “Some theorems on sorting”. *SIAM J. on Applied Mathematics* 17, 1969, 1-6.
- [6] Raz, T., Herer, Y. T. and Grosfeld-Nir, A. “Economic optimization of off-line inspection”. *IIE Trans* 32, 2000, 205–217.
- [7] Sheu, S. H., Chen, Y. C., Wang, W. Y., and Shin, N. H. “Economic optimization of off-line inspection with inspection errors”. *Journal of the Operational Research Society* 54, 2003, 888–895
- [8] Wang, C-H. “Economic off-line quality control strategy with two types of inspection errors”. *European Journal of Operation Research* 179, 2007, 132–147.

DOES THE SYSTEMS THEORY APPROACH APPLY TO SERVICES?

Adenekan (Nick) Dedeke, University of Massachusetts, Boston, (617)-287-7888,
adenekan.dedeke@umb.edu

ABSTRACT

This paper presents the use of systems theory approach to develop a definition for services, a conceptual model for service systems and a services positioning matrix. We propose a two-dimensional services matrix that creates four types of services systems. The matrix enables services researchers to integrate systems coupling and service acts in the framing of services. The systems theory approach also allows us to frame services operations in a more comprehensive manner.

Keywords: services matrix, systems theory, services

Introduction

The services industry segment of most western countries has grown significantly in the past two decades [12]. U.S. labor statistics reveal that over 80% of the U.S. economy in 2006 was accounted for by the services sector [19]. The growth created at least two effects. First, the contribution of services products to total corporate profits of many manufacturing firms, such as General Motors, General Electric and of services organizations, have increased [13]. This suggests that, in the next decades, the level and pace of economic growth, as measured by gross domestic product, and the profitability of both manufacturing and services firms, would be determined by the degree of productivity that is achieved in the services sector. Unfortunately, the research of the field of services has not kept pace with this new reality [10, 18]. The purpose of this paper is to suggest ways in which systems theory approach could help to understand services operations in a more systematic manner.

Value of a Systems Theory Approach for Services

Systems theory refers to the concepts of science of systems that resulted from Bertalanffy's General System Theory (GST). His purpose was to find a unifying theory for the whole field of science in general [21]. Through the extension of the theory by researchers and its adaptation for scientific inquiry, it has become a multidisciplinary and multiperspectival field. The assumptions of the systems theory approach make it especially applicable to the services field. First, in the current state, the field of services research seems to be in need of the type of unification that systems theory can provide [10, 14]. Second, systems theory approach is grounded on the premise that the behavior of a system is more clearly understood by an inquiry into the interrelationships between elements of the system than through the study of the elements in isolation. According to the approach, the focus of researchers would shift from elements themselves to the structure between the elements of a system and to the interactions between the elements [1, 8]. This view is meaningful for most service systems, where the relationships between the elements and the interactions between them during co-production determine the output of the services system. The centrality of interactions is one of the reasons why the customer contact-based services classification [2] became a popular framework for framing services encounters. A weakness of the customer-contact based view was that it focused primarily on the customer's contact and interaction with a system. A systems theory approach may enable us to include a broader set of interactions as one

defines service encounters. This may also permit researchers to propose and test richer hypotheses that go beyond the ones that relate the degree of customer contact to the potential productivity of a system.

Definition of Services Systems

The challenge of developing a scientific approach to services ought to include the definition of the discipline. Metters and Maruchek [10] confirm the lack of consensus in the definition of services and in the classification of services operation in extant literature. For example, there are nominal definitions of services, such as whatever is not manufacturing is services [14]. While this definition might be useful for reporting statistics about an economy, it is somewhat vague for scientific research. Academic literature provides at least three kinds of definition for services. First, there are definitions that stem from the economics discipline and research. Hill [6] defines services as a change in condition of a person or a good belonging to some economic entity, brought about as the result of the activity of some other economic entity, with the approval of the first person or economic entity. Second, there are definitions of services from the operations management field that emphasize the degree of contact during the service act. According to this view, services can be viewed as systems that enable the creation of a service product under varying degrees of customer contact with the service system [2]. The limitation of this definition is its exclusion of services systems in which customers do not directly interact with the system during the service creation process [5]. Samson and Froehle [14] modify the definition by introducing the concept of customer inputs into it. According to Samson and Froehle, the presence of customer inputs, not of the customer contact, is a necessary and sufficient condition that differentiates production processes from service processes. Though Samson and Froehle may have succeeded in overcoming one critique of the Chase definition, the one that they propose does not provide direction or to answer some critical issues of services. While their definition identifies customer inputs as the necessary condition that distinguishes production from services processes, it does not define the factors that differentiate services from each other. Furthermore, it does not attempt to define various kinds of output that may constitute services delivery. Finally, the definition is silent on the involvement of workers and assets such as information technology and similar resources, during the service encounter. These are questions that we believe a systems theory based view of services would help to isolate. We believe that a systems theory oriented approach could help develop a broader services definition that would enable one to differentiate services types.

In this section, we use a systems theory based methodology to develop a definition for a services system. Our methodology, based on Checkland's work [3], involves the development of a root definition for services and thereafter the development a conceptual model that is based on the root definition. For a detailed introduction of the procedures involved in the creation of root definitions, we refer the reader to Checkland [3]. Using this procedure, we propose the following definition for services systems:

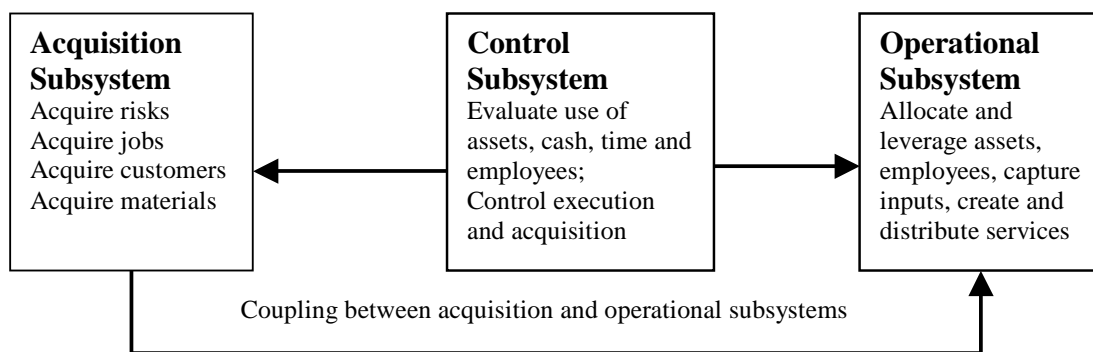
A services system, is a privately owned and profit-oriented enterprise which, given the intensity of competition in its business environment, allocates its staff and acquired resources to clients and leverages them to achieve negotiated results and relationships, and to shoulder negotiated responsibilities, and risks for and on behalf of its customers, at prices which are sufficiently above the firm's services provision costs. Services employees are attracted, hired, and allocated to ensure that customer inputs are interpreted correctly, that resources are operational, and that customer-interactions with the services system occur with minimum glitch.

This definition is broader than most definitions that have been proposed in literature. First, its focus is not to define services by differentiating production from services systems, rather it defines services.

Second, it does not attempt to develop a definition that fits preconceived characteristics of services systems; rather its focus is on specifying the minimal number of elements that constitutes a services system. The definition introduces the profit-orientation as the economic basis of the services system. It also identifies private ownership as its holding structure. The competitive environment, which is constituted by firms providing similar, substitutes and competing services, are identified as factors which impact services firms. The definition proposes three minimal functions for a services system, including the acquisition of resources and customers. Furthermore, the leveraging of resources is identified as a key transformation process of a services firm. The root definition identifies two types of service outcomes. These are the negotiated results, such as tangible physical product changes, and intangible outcomes, such responsibilities, relationships, and risks.

Based on the root definition for services, we proceed to derive a conceptual model for services system Figure 1 shows the minimum set of subsystems that one would expect based on the root definition, provided consists of three elements: the acquisition subsystem, the operational subsystem and the control subsystem. The purpose of the control system is to monitor the manner in which the acquisition and the operational subsystems are operating. However, the control system will also responsible for generating feedback to each of the systems, which would be used to improve performance. The functions of the acquisition and the operational subsystems are evident from figure 1.

Figure 1: Conceptual model of a services system



We propose that one of the ways in which services differ is the manner in which the subsystems are coupled together. For example, the acquisition and operational subsystems could be tightly or loosely coupled. We will use this attribute in the next section when we develop a services matrix.

Classifying Services by the Degree of System Coupling

The classification of services systems has been an area of research that has received a lot of attention from researchers in literature. Initially, researchers attempted to develop a matrix that would replicate the product-process matrix that was developed by Hayes and Wheelwright [6] for the manufacturing segment, for the services sector. The Hayes and Wheelwright matrix facilitates the choice of a production system once the product volume and the product variety characteristics were established. Most of these efforts to replicate the Hayes et al for services were not successful. Hence, researchers shifted their emphasis on the development of frameworks that were intended to help classify services firms or services processes according to the characteristics that they exhibit [4]. One premise of these classification schemes was that different services characteristics require different management strategies [15, 17]. Several two dimensional classification schemes have been introduced in literature. Schmenner's matrix [15] differentiated services by degree of labor intensity and degree of customer

contact. Shostack [16] proposed a classification that was based on complexity and divergence. Wemmerloev [24] suggested the dimensions of customer contact and the degree of rigidity/fluidity of processes. Kellog and Nie [24] suggested a classification scheme based on customer influence on process and the service package structure (degree of customization). Collier and Meyer [14] used the dimensions of number of customer routes built into a services system and the repeatability of services encounter. We list only two dimensional classification schemes because they constituted the majority of the proposals that we found in literature. Furthermore, we list them because our paper proposes a two dimensional scheme.

The need for a new unifying classification scheme has been echoed in literature. Verma [20] found that the classification scheme that was proposed by Schmenner [15] failed to empirically differentiate services from each other. Collier and Meyer [4] found weak or no empirical support for the clarity of the constructs that were presented by Kellog and Nie [7], only for the constructs that they proposed. They call for positioning frameworks with better criteria, clearer causation directions, and improved empirical measures.

Figure 2: Service positioning matrix

Systems coupling

<i>Loosely coupled</i>	<i>Competence sharing factories, e.g. research and innovation networks [2]</i>	<i>Mass processing services, e.g. global supply chain partnerships, wireless nets [3]</i>
<i>Tightly coupled</i>	<i>Service factories, e.g. medical practices, schools, car wash, hospitals, restaurants [1]</i>	<i>Mass access (membership) services, e.g. computing and phone, gas and water nets [4]</i>

Service acts types

Few/mostly unique Many/mostly standard

In their analysis of the existing frameworks, Samson and Froehle [14] stated that a general critique of the frameworks was the lack of criteria, which determine whether a system is standardized or predictable. We believe that the scheme that we present meets many of these unfulfilled expectations. Figure 2 shows the services positioning matrix that we propose. We classify the services using two dimensions; by service acts provided to customers and by the degree of coupling between the elements that constitute the services system. We differentiate two kinds of service acts, namely, mostly standardized and mostly unique ones. Our systems coupling dimension differentiates between tightly and loosely coupled systems. Services acts are defined as being mostly standardized when only routine knowledge or skills are required by clients and employees to deliver services outputs. Otherwise, the services acts are defined as being unique. Services coupling dimension identifies tightly coupled and loosely coupled systems. Tightly coupled services systems are defined as those that require that presence of employees, clients, and service targets to be concurrently present in the same place at the same time for service creation to occur, or systems that require a real-time physical connection between clients and a provider's network during services co-production and consumption. Loosely coupled systems are defined as those which do not require concurrent presence of employees, clients, service targets at the same place and time for service creation to occur, or those which do not require real-time physical

connection between clients and a service provider's network during service consumption. Figure 2 suggests that when customers require mostly unique services, the services system design has to be focused on the accomplishment of results. However, when the services co-produced are mostly standardized, services system design would emphasize the optimization of services processes.

Organization science literature already has theories that are linked to the coupling of systems [9]. Miller [11] suggests that loose coupling of systems might be a necessity that accompanies growth. He postulates that as a system grows and adds more components, the components in general become increasingly independent in decision-making. Weick [22, 23] postulates that loosely coupled systems are flexible and better capable of improvising and self-design. These hypotheses offer us the opportunity to test their validity for services organizations.

Conclusion

This paper presents the use of systems theory approach to develop a definition for services, a conceptual model for service systems and a services positioning matrix. Future work would extend this work by identifying measures that could be used to define the degree of coupling of services firms. The matrix also needs to be compared and contrasted to theoretical approaches that have been suggested in literature. Overall, we believe that the systems theory approach is a powerful approach that is valuable for developing a systems theory based theory of services systems.

REFERENCES

- [1] Bailey, K. D. *Sociology and the New Systems Theory: Toward a theoretical synthesis*. New York, NY: State of New York Press, 1994.
- [2] Chase, R. B. "The customer contact approach to services: Theoretical bases and practical extensions," *Operations Research*, 1981, 29(4), 698-706.
- [3] Checkland, P. *Systems Thinking, Systems Practice*. New York, NY: John Wiley & Sons, 1999.
- [4] Collier, D. A. and Meyer, S. M. "An empirical comparison of service matrices," *International Journal of Operations & Production Management*, 2000, 20(6), 705-729.
- [5] Froehle, C. M. and Roth, A. V. "New measurement scales for evaluating perceptions of the technology-mediated customer service experience," *Journal of Operations Management*, 2004, 22(1), 1- 21.
- [6] Hayes, R. H. and S. C. Wheelwright, S. C. "Link manufacturing process and product life cycles," *Harvard Business Review*, 1979, 57(1), 133-140.
- [7] Kellog, D. L. and Nie, W. "Framework for strategic service management," *Journal of Operations Management*, 1995, 13(4), 323-337.
- [8] László, E. *The Systems View of the World*. New York, NY: George Braziller, 1972.

- [9] Lawrence, P. R. and J. W. Lorsch, J. W. *Organization and environment*. Homewood, Ill: Irwin, 1969.
- [10] Metters, R. and Marucheck, A. "Service management – Academic issues and scholarly reflections from operations management researchers," *Decision Sciences*, 2007, 38(2), 195-214.
- [11] Miller, J. G. *Living Systems*. New York, NY: McGraw-Hill, 1978.
- [12] OECD. *The OECD Science Technology and Industry Outlook*. Paris: OECD, 2004.
- [13] Quinn, J. B. *Intelligent Enterprise: A knowledge and service based paradigm for industry*. New York, New York: The Free Press/McMillan Inc., 1992.
- [14] Samson, S. E. and Froehle, C. M. "Foundations and implications of a proposed unified services theory," *Production and Operations Management*, 2007, 15(2), 329-343.
- [15] Schmenner, R. W. "How can service businesses survive and prosper?," *Sloan Management Review*, 1986, 27(3), 21-32.
- [16] Shostack, G. L. "Service positioning through structural change," *Journal of Marketing*, 1987, 51(1), 34-43.
- [17] Silvestro, R. "Positioning services along the volume-variety diagonal," *International Journal of Operations & Production Management*, 1999, 19(3/4), 399-420.
- [18] Smith-Daniels, V. "The next frontier in services research," *Decision Sciences*, 2007. 38(2), 187-192.
- [19] U.S. Department of Labor, "The employment situation, Bureau of Labor Statistics," available at: <http://www.bls.gov/news.release/pdf/empsit.pdf>, 2006.
- [20] Verma, R. "An empirical analysis o management challenges in service factories, service shops, mass services and professional services," *International Journal of Service Industry Management*, 2000, 11(1), 8-25.
- [21] von Bertalanffy, L. *General System Theory: Foundations, Development, Applications*, New York, NY: George Braziller, 1968.
- [22] Weick, K. *Making Sense of the Organization*, Malden, MA: Blackwell Publishers Ltd., 2001.
- [23] Weick, K. E. "Organization design: Organizations as self-designing systems," *Organizational Dynamics*, 1997, 6(2), 30-46.
- [24] Wemmerloev, U. "A taxonomy for service processes and its implication for systems design," *International Journal of Service Industry Management*, 1990, 1(3), 13-27.

MULTICRITERIA DECISION-MAKING IN THE ETHANOL PRODUCTION PROBLEMS: A GOAL PROGRAMMING APPROACH

Kenneth D. Lawrence, School of Management, New Jersey Institute of Technology, Newark, NJ 07102, carpetfour@yahoo.com, 1-973-596-6425

Dinesh R.Pai, Rutgers Business School, Rutgers University, Newark, NJ 07102, dineshp@rutgers.edu, 1-973-353-5671

Ronald Klimberg, Decision & System Sciences Department, Haub School of Business, Saint Joseph's University, 5600 City Avenue, Philadelphia, PA 19131, klimberg@sju.edu, 1-610-660-1625

Sheila M. Lawrence, Department of MSIS, Rutgers University, Piscataway, NJ 08854, lawrencem5@aol.com, 1-973-596-6425

ABSTRACT

This paper deals with a supply chain problem involving the production of ethanol and various by-products. The main objective of this problem is to minimize the total production costs within a maximum allowable budget without violating the demand constraints. The secondary objectives are adhering to: given pollution norms and waste minimization constraints. A preemptive goal programming model is developed to solve the ethanol production planning problem. **Keywords:** Multicriteria, Goal programming, Ethanol

INTRODUCTION

Rising fuel prices and policy initiatives have continued to stimulate renewable fuels including ethanol. Since, the mid-1990s the number of ethanol facilities and the plant size has increased gradually. By mid-2006, nearly 100 ethanol facilities in the United States were producing more than four billion gallons of ethanol annually, with 50 to 100 million gallon per year plants as standard size [3]. This paper deals with a supply chain problem involving the production of ethanol and various by-products. The process includes material sources, the processing mills and the customers. The primary objectives are to minimize the cost of: source materials, production (wet or dry milling), and transportation of final products. Additionally, the minimization of pollution at milling sites is another important management objective.

The goal programming (GP) technique provides an analytical framework that a decision-maker can use to provide optimal solutions to multicriteria and conflicting objectives. The GP and its variants have been applied to wide variety of problems [4] [6]. The use of GP in process industry problems is not new. Krajnc [4] investigates possibilities of attaining zero-waste emissions in case of sugar production. Arthur and Lawrence [1] designed a GP model to develop production and shipping patterns for chemical and pharmaceutical industries.

The model presented in this paper is designed to illustrate how preemptive GP can be used as an aid to solving multicriteria production related problems. Our ultimate goal is to develop a max-min model as well as a fuzzy goal programming model. Zimmermann [7] proposed the first

method for solving fuzzy linear programming problems. Fuzzy optimization has been focusing, first in solving models which reflect real life uncertainty, and second on transforming them into equivalent crisp problems to benefit from efficient existing solving algorithms. Fuzzy decision is a combination of goals and constraints, denoting a max-min model because it considers that the best fuzzy decision is the union of the aggregated intersections of goal and constraints [2].

FORMULATION

In order to formulate the model, we define the following:

Notations

Indices:

- i : the index of the sources of corn
- j : the index of the production process
- k : the index of the product (ethanol, corn oil, dry meal, corn gluten, live stock feeding, waste)
- l : the index of the customer groups

Sets:

- I : the set of the sources of corn
- J : the set of the production process
- K : the set of the products
- L : the set of the customer groups

Parameters

- d_{kl} = Monthly demand for product type k for customer group l .
- C_{ijkl} = Total cost of producing and shipping the k^{th} product type (including raw materials) from the i^{th} corn source through the j^{th} milling location to the l^{th} customer group.
- t_{ijk} = Unit time to produce a unit of the k^{th} output type at the j^{th} mill.
- b_j = Production capacity at the j^{th} mill.
- p_{jk} = Pollution level at the j^{th} mill for the production of a unit of the k^{th} output type (Gallons of water)
- H_j = Number of hours available on a yearly basis for the j^{th} mill
- TC = Total cost
- PG_j = Pollution limit for the j^{th} mill
- TW = Total waste generated
- P_q = Priority labels, where, $q = 1, 2, 3$

Variables

X_{ijkl} = the amount of product type k produced from corn from source i in mill type j for customer group l .

d_1^+ = the deviation variable of overachievement of the goal 1

d_1^- = the deviation variable of underachievement of the goal 1

d_2^+ = the deviation variable of overachievement of the goal 2

d_2^- = the deviation variable of underachievement of the goal 2

d_3^+ = the deviation variable of overachievement of the goal 3

d_3^- = the deviation variable of underachievement of the goal 3

Goal constraints and objective functions

Goal 1: Minimize total cost

$$Z_1 = \sum_{l=1}^L \sum_{k=1}^K \sum_{j=1}^J \sum_{i=1}^I C_{ijkl} X_{ijkl} - d_1^+ + d_1^- \leq TC, X_{ijkl} \geq 0 \quad (1)$$

Goal 2: Reduction in level of pollution for the j^{th} mill

$$Z_2 = \sum_{k=1}^K \sum_{j=1}^J p_{jk} X_{ijkl} - d_2^+ + d_2^- \leq PG_j \quad \forall i = 1, \dots, I, \forall j = 1, \dots, J \quad (2)$$

Goal 3: Reduction in level of waste

$$Z_3 = \sum_{i=1}^I \sum_{j=1}^J \sum_{l=1}^L X_{ijkl} - d_3^+ + d_3^- \leq TW \quad \text{for } k = 6 \quad (3)$$

The objective function of the model is to minimize the deviation variables corresponding to various goals. Highest priority is assigned to the total cost goal, to minimize the total cost. Therefore, the undesirable deviational variable in the first priority is overachievement of goal 1, i.e. d_1^+ should be minimized. Then, the first priority function in the objective function is $P_1 d_1^+$. In the second priority of pollution level, we wish to minimize the pollution level from each mill to a predetermined safe limit. Therefore, the undesirable deviational variable in the second priority goal is overachievement of the goal 2 i.e. d_2^+ should be minimized. Then, the second priority function in the objective function is $P_2 d_2^+$. Finally, in the third priority of waste level, we wish to restrict the waste produced within a predetermined limit. Therefore, the undesirable deviational variable in the third priority goal is overachievement of the goal 3 i.e. d_3^+ should be minimized. Then, the third priority function in the objective function is $P_3 d_3^+$.

Hence the objective function for the goal programming model is:

$$\text{Minimize: } Z = P_1 d_1^+ + P_2 d_2^+ + P_3 d_3^+ \quad (4)$$

Constraints

The objective functions formulated in the previous section are restricted by two sets of constraints. They are the demand constraints, and the time constraints.

$$\sum_{k=1}^K \sum_{l=1}^L X_{ijkl} \geq d_{kl}, \quad \forall i = 1, \dots, I, \forall j = 1, \dots, J \quad (5)$$

$$\sum_{l=1}^L \sum_{k=1}^K \sum_{i=1}^I t_{ijk} X_{ijkl} \leq H_j \quad \forall j = 1, \dots, J \quad (6)$$

$$X_{ijkl}, d_1^+, d_1^-, d_2^+, d_2^-, d_3^+, d_3^- \geq 0, i \in I, j \in J, k \in K, l \in L \quad (7)$$

Constraint (5) ensures that the customer demands are met, whereas, the constraint (6) limits the hours available for processing on each type of milling. Constraint (7) ensures that all the decision variables are non-negative.

REFERENCES

- [1] Arthur, J., and Lawrence, K (1982) Multiple goal production and logistics planning in a chemical and pharmaceutical company, *Computers & Operations Research*, 9(2), 127–137.
- [2] Bellman, R., and Zadeh, L (1970) Decision-Making in a Fuzzy Environment. *Management Science*, 17(4), 141-164.
- [3] EAA (2006), *Economic Impacts of Ethanol Production, Ethanol Across America, USA*.
- [4] Ignizio, J. (1976), *Goal Programming and Extensions*, Massachusetts, Lexington Books.
- [5] Krajnc, D., Mele, M., and Glavic, P (2007), Improving the economic and environmental performances of the beet sugar industry in Slovenia: increasing fuel efficiency and using by-products for ethanol, *Journal of Cleaner Production*, 15, 1240-1252.
- [6] Romero, C. (1991), *Handbook of Critical Issues in Goal Programming*, Oxford, Pergamon Press.
- [7] Zimmermann, H (1978), Fuzzy Programming and Linear Programming with Several Objective Functions. *Fuzzy Sets and Systems*, 1, 45-55.

A MULTIPLE OBJECTIVE GOAL PROGRAMMING APPROACH TO SINGLE-LINE, MULTI-PRODUCT DAIRY PRODUCTION SCHEDULING

By Virginia M. Miori, Ph.D. St Joseph's University Philadelphia, PA

ABSTRACT

The application of optimal methods for production scheduling in the dairy industry has been limited. Within supply chain terminology, dairy production was generally considered a push process but with advancements in automation, the industry is slowly transforming to a pull process. This paper presents triplet notation applied to the production scheduling of a single production line used for milk, juice, and carnival drinks. Once production and cleaning cycles are characterized as triplets, the problem is formulated with multiple objectives, paving the way for a solution employing goal programming and tabu search techniques.

INTRODUCTION

Milk is initially produced in two variations, skim, and whole. One percent and two percent milk are the result of blending appropriate ratios of skim and whole milk. Therefore, a single filling line may simultaneously produce all four varieties of milk [5]. Other product families such as orange juice, carnival drinks, and buttermilk are also processed using the same filling line although separate storage tanks are used for each product family.

Product is piped into one of two bowls for filling. Between batches of these different product families, the equipment must be cleaned. Less extended cleaning can also be required between products within the same product family.

A primary goal of this scheduling approach is to facilitate the transition to a pull system. Order due dates will act as constraints in the problem formulation and ultimately guide the production order. Customers have standing due dates each week for their orders and these dates must be met. Additionally the scheduling approach will reduce inventory held. In the past excess inventory was carried when it was uncertain as to whether an order could be slotted into the production schedule.

DAIRY OPERATION

Milk, orange juice, carnival drinks, and buttermilk are all held in their own storage tanks. Product is transferred from the storage tanks into two filling bowls. In the case of milk, one bowl is filled with

whole milk and one with skim milk and all varieties of milk can be produced through adjusting the ratios of whole and skim milk used. All other products are made off line and placed in a single filling bowl. Note that variations in product type results in a need for cleaning the filling bowls, supply lines, and valves between production runs. The bottles are labeled and then filled, each label not only stating the contents of the bottle, but the specific customer name.

All products can be generalized into product families and the necessary cleaning cycles are determined based on both product family and purity of product.

Product Families

Though orders are placed for individual products, production times and cleaning times are determined by product family. Table 1 provides a sample of product families and individual products. This is certainly not an exhaustive list, but it is useful in illustrating production and cleaning sequences.

Product Family	Product Subtypes
Milk (M)	Whole Milk 2 % Milk 1% Milk Skim Milk
Orange Juice (OJ)	No Pulp OJ Medium Pulp OJ High Pulp OJ Calcium Fortified OJ
Carnival Drink (CD)	Orange Drink Fruit Punch Grape Drink
Buttermilk (B)	Whole Buttermilk Low fat Buttermilk

Table 1: Product Families

TRIPLETS

Cleaning Transitions

A full cleaning cycle (FC) is required between consecutive production runs of any two product families. Cleaning may also be required between production runs of individual products within a product family. For example, a rinse (R) would be required after high pulp OJ if no pulp OJ were the next product to be produced. A rinse (R) would not be required however, if the production were to proceed in the opposite order. In any sequential production runs between which cleaning is not required, a cleaning type of no-clean (NC) is designated with a completion time of zero (0) minutes. The required cleaning cycles are presented in Table 2.

Packaging Transitions

Non-milk products are produced and sold in single gallon quantities. The milk products however, may be packaged in three different ways: single gallon, two-pack of gallons, or four-pack of gallons with each pallets holding only one package type. We must not only account for cleaning between products, but we must also account for any package transitions. The production sequence is altered when shifting between single and multi-packs, multi-packs and singles and within multi-packs. A bundler is used to combine single gallons into multi-packs. Single gallons are diverted from the bundler while gallons to be combined into multi-packs are sent to the bundler.

		Milk				OJ				CD			B	
		W	2	1	S	NP	MP	HP	C	O	F	G	L	W
Milk	W	NC	R	R	R	FC				FC			FC	
	2	NC	NC	R	R									
	1	NC	NC	NC	R									
	S	NC	NC	NC	NC									
OJ	NP	FC				NC	NC	NC	NC	FC			FC	
	MP					R	NC	NC	R					
	HP					R	R	NC	R					
	C					R	R	R	NC					
CD	O	FC				FC				NC	R	R	FC	
	F									R	NC	R		
	G									R	R	NC		
B	L	FC				FC				FC			NC	NC
	W												R	NC

Table 2: Required Cleaning Cycles

Introduction of Triplets

The triplets used in dairy production scheduling are similar to the triplets applied to transportation routing by Miori [7]. A scheduling triplet is composed of a productive activity (first node), an unproductive activity (second node) and followed by a subsequent productive activity (third node). For dairy production, this means milk production, followed by transition between the first node production run and the third node production run. The third node acts only as a placeholder. As triplets are chained together to create a schedule, the third node of a triplet must match the first node in the subsequent triplet. This creates a logical connection between the production activities.

Production and transition times are determined for each node of the triplet. The triplets are generated from the orders for products so that any triplet corresponds to a single order. Multiple triplets are created for every order to allow more sequencing options to be considered, though once the order has been produced, all associated triplets are eliminated from consideration.

The triplet time is the sum of the individual node times. The third node is always considered to have a time of zero (0). The production time reflects a base production rate of 5 minutes per pallet. The transition time includes the accumulation of cleaning and packaging transitions necessary between the first and third nodes. Recall that the cleaning transition times are deterministic and for purposes of this paper, the stochastic packaging transitions are treated as deterministic. Costs are accumulated in the same manner as times.

SCHEDULING MODEL

The use of triplets has been established in transportation modeling where the first leg of the triplet designates a loaded movement and the second leg designates an empty movement. We carry this idea over to production scheduling where the first leg of the triplet represents a production run and the second leg represents a transition which includes a cleaning cycle and a packaging transition. We represent the triplets with a (0,1) decision variable $x_{tuv}^{p(o),c}$ and cost c_{pc} (t = production start time, u = production end time coincident with cleaning start time, v = cleaning end time, $p(o)$ = product demanded in order o , and c = transition cycle type).

Demand and order delivery dates are known in advance and completed orders must be staged for transport by their load-out date and time. The scheduling horizon is 6 days long, with two shifts per day. At the end of the second shift each day, the filling line is emptied and cleaned so that no production runs carry over to the following day. The production-scheduling objective(1) is to minimize production cost that implicitly minimizes machine downtime:

$$\text{Min} \sum_{p(o),ctuv} c_{p(o),c} x_{tuv}^{p(o),c} . \quad (1)$$

The initial constraint ensures that all demand scheduled for immediate receipt by customers be served (2). If production time is available, the next day's demand may begin processing before close of the current day's production. Demand is denoted as $y_{p(o),d}$ (d = day order due) where the y 's are integer valued parameters.

$$\sum_{ctuv} y_{p(o),d} x_{tuv}^{p(o),c} \geq 1 \quad \forall y_{p(o),d} \geq 1 \quad (2)$$

As in the traveling salesman problem, conservation of flow (3) must also be achieved.

$$\sum_{p(o),ctuxy} (x_{tuv}^{p(o),c} - x_{vxy}^{p(o),c}) = 0 \quad \forall v \quad (3)$$

Schedule feasibility must be maintained. We accomplish this through constraint (4), which maintains successive process times and constraint (5), which forces on-time order fulfillment.

$$t \leq u \leq v \quad \forall x_{tuv}^{p(o),c} \quad (4)$$

$$u \leq D_{p(o)} \quad \forall x_{tuv}^{p(o),c} \quad (5)$$

Finally, the decision variables are all binary (6).

$$x_{tuv}^{p(o),c} \in (0,1) \quad (6)$$

Multiple Objective Formulation

In order to take advantage of the capabilities of goal programming and tabu search techniques, the dairy model must first be formulated at a multiple objective model. Minimization of production costs is joined by other objectives such as production line sequencing, fulfilling load-out time window requirements and service of all demand.

Before we discuss the specific objectives, we note that the solution approach we will be taking is that of goal programming. Therefore, each objective must have a target value. We then define the positive and negative deviations from this target to be used in the goal programming minimization.

Goal programming allows for two approaches, one in which each objective is weighted and the second in which the objectives are prioritized and the lexicographic minimum is found. We use the second approach since there is a clear priority structure inherent in our multiple objectives. The complete multiple objective formulation is:

$$\text{Minimize } z_1 = \sum_{ij} d_{1ij}^- \quad (7)$$

$$\text{Minimize } z_2 = \sum_{ij} d_{ij}^+ \quad (8)$$

$$\text{Minimize } z_3 = \sum_v (w_{3v}^+ d_{3v}^+ + w_{3v}^- d_{3v}^-) \quad (9)$$

$$\text{subject to: } \sum_{vk} y_{p(o),d} x_{tuv}^{p(o),c} - d_{1ij}^- = 1 \quad \forall y_{ij} = 1 \quad (10)$$

$$\sum_{ijkv} c_{p(o),c} x_{tuv}^{p(o),c} - d_{2ij}^+ = 0 \quad (11)$$

$$\sum_{ijlmv} (x_{ijk}^v - x_{klm}^v) = 0 \quad \forall k \quad (12)$$

$$t \leq u \leq v \quad \forall x_{tuv}^{p(o),c} \quad (13)$$

$$y_{ij} \in (0,1) \quad x_{ijk}^v \in (0,1)$$

Though the cost minimization was the single objective in our formulation, it will not be the highest priority objective in the MO formulation. The highest priority is assigned to the minimization of unserved demand (7). Each unit of demand is identifiable as the first leg of a triplet. We wish to accumulate the triplets assigned to the schedule and minimize the difference between this sum and the total units demand in the particular instance of the problem.

The constraints tied to this objective (10) represent the sum of all demand served and the negative deviation from the goal of serving all demand. Recall from the previous section that the integer parameters $y_{p(o),d}$ provide a representation of the demand for each lane. The positive deviation is not considered since we would be unable to service more demand than was available.

Meeting production time windows is the second priority objective. The goal is the availability of completed milk production at particular time. The deviations from this times are weighted; late order production is unacceptable and therefore any negative deviation is heavily weighted to discourage this. All deviations have a negative weight.

The cost minimization objective (8) falls next in priority order and has a target of zero. We now eliminate the negative deviation from consideration due to the restriction that all costs must be greater than or equal to zero. This approach leaves us with a single deviation measure. The associated constraints (11) follow. The cost for each production run/transition pair is accumulated and the positive deviation determined between that cost and the target of zero cost.

We are left to define the remaining constraints of the formulation. These include the conservation of flow constraints (13) and the route sequencing constraints (13). We may draw these directly from our single objective formulation.

CONCLUSION

The dairy production scheduling model lends itself very nicely to a multiple objective formulation. The traditional traveling salesman conservation of flow constraint remains in the multiple objective formulation while the remaining constraints positioned as additional objectives. We are now prepared to implement a goal programming solution method using tabu search.

References available upon request from Virginia Miori.

USING DATA ENVELOPMENT ANALYSIS (DEA) TO FORECAST BANK PERFORMANCE

Ronald K. Klimberg, Saint Joseph's University, Haub School of Business, Philadelphia, PA, 19131, 610-660-1625, klimberg@sju.edu

Kenneth D. Lawrence, School of Management, NJIT, Newark, NJ 07102, (973) 596-6425, carpetfour@yahoo.com

Tanya Lal., Saint Joseph's University, Haub School of Business, Philadelphia, PA, 19131, 610-660-1625, t1361386@sju.edu

ABSTRACT

Forecasting is an important tool used by businesses to plan and evaluate their operations. One of the most commonly used techniques for forecasting is regression analysis. Often forecasts are produced for a set of comparable units which could be individuals, groups, departments or companies that perform similar activities such as a set of banks, a group of managers and so on. We apply a methodology that includes a new variable, the comparable unit's DEA relative efficiency, into the regression analysis. The results of applying this methodology to the performance of commercial banks will be presented.

(Keywords: Forecasting, Data Envelopment Analysis, Regression)

INTRODUCTION

Quantitative forecasting models, even rather sophisticated models, are easier to develop and use today as result of our improving computer technology. These quantitative forecasting techniques use historical data to predict the future. Most quantitative forecasting techniques can be categorized into either time series approaches or causal models. Time series forecasting techniques are forecasting techniques that only use the time series data itself and not any other data to build the forecasting models. These time series approaches isolate and measure the impact of the trend, seasonal, and cyclical time series components. Causal models use a set of predictor/independent variables, possibly also including the time series components, that are believe to influence the forecasted variable. One of the most popular causal model approach is regression analysis. Regression techniques employ the statistical method of least squares to establish a statistical relationship between the forecasted variable and the set of predictor/independent variables.

Many forecasting situations involve producing forecasts for comparable units. A comparable unit could be an individual, group of individuals, a department, a company, and so on. Each comparable unit should be performing similar set of tasks. When applying regression analysis, the established statistical relationship is an average relationship using one set of weights assigned to the predictor/independent variables. However, when regression is applied to a set of comparable units the relative weight/importance of each of the predictor/independent variables

will most likely vary from comparable unit to comparable unit. For example, if advertising is an independent variable, one comparable unit might emphasize advertising more (or less) than other comparable units. Either way is not necessarily better nor worse, it is just how that particular comparable unit emphasizes advertising. As a result, in some cases, the regression model could provide forecast estimates that are too high or too low.

In this paper, we will apply and extend some of our recent previous work, Klimberg et al., [3, 4], in which we introduced a methodology that incorporates into the regression forecasting analysis a new variable that captures the unique weighting of each comparable unit. This new variable is the relative efficiency of each comparable unit that is generated by a non-parametric technique called data envelopment analysis (DEA). In the next section, we provide a brief introduction to DEA. Subsequently, we discuss the methodology and present the results of applying our methodology to a data set of commercial banks. Finally, the conclusions and future extensions are discussed.

DATA ENVELOPMENT ANALYSIS (DEA)

DEA utilizes linear programming to produce measures of the relative efficiency of comparable units that employ multiple inputs and outputs. DEA takes into account multiple inputs and outputs to produce a single aggregate measure of relative efficiency for each comparable unit. 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. The Charnes, Cooper and Rhodes (CCR) DEA model [1] is a linear program that compares the ratio of weighted outputs to weighed inputs, i.e., efficiency, for each comparable unit. The efficiency of the k^{th} comparable unit (i.e., E_k) is obtained by solving the following linear formulation:

$$\text{MAX } E_k = \sum_{r=1}^t u_r Y_{rk}$$

s.t.

$$\sum_{i=1}^m v_i X_{ik} = 1$$

$$\sum_{r=1}^t u_r Y_{rj} - \sum_{i=1}^m v_i X_{ij} \leq 0 \quad j=1, \dots, n$$

$$u_r, v_i \geq \epsilon \quad \forall r, i$$

where:

Parameters

Y_{rj} = amount of the r th output for the j th comparable unit;

X_{ij} = amount of the i th input for the j th comparable unit;

t = the number of outputs;

m = the number of inputs, and;

n = the number of comparable units;

ϵ = is a small infinitesimal value;

Decision Variables

u_r = the weight assigned to the r th output, and;

v_i = the weight assigned to the i th input.

The CCR DEA formulation determines objectively the set of weights, u_r and v_i , that maximizes the efficiency of the k th comparable unit, E_k . The constraints require the efficiency of each comparable unit, including the k th comparable unit, not to exceed 1, and the weights, u_r and v_i , must be positive. A similar DEA formulation must be solved for each comparable unit. A comparable unit is considered relatively inefficient (i.e., $E_k < 1$) if it is possible to increase its outputs without increasing inputs or decrease its inputs without decreasing outputs. A comparable unit identified as being efficient (i.e., $E_k = 1$) does not necessarily imply absolute efficiency. It is only relatively efficient as compared to the other comparable units that are being considered. These efficiency ratings allow decision-makers to identify which comparable units are in need of improvement and to what degree.

Each efficiency score measures the relative efficiency of the comparable unit. These efficiency scores can be used to evaluate performance of the comparable units and provide benchmarks. Nevertheless, besides each efficiency score being comprised of a different set of inputs and outputs values, each comparable unit's efficiency score includes a unique set of weights. The DEA process attempts to find objectively the set of weights which will maximize a comparable unit's efficiency. Therefore, the DEA model has selected the best possible set of weights for each comparable unit. The variation of these weights from comparable unit to comparable unit allows each comparable unit to have their own unique freedom to emphasize the importance of each of these input and output variables in their own way. How well they do this is measured by the efficiency score.

Since the Charnes, et al.'s 1978 paper, there have been thousands of theoretical contributions and practical applications in various fields using DEA. DEA has been applied to many diverse areas such as: health care, military operations, criminal courts, university departments, banks, electric utilities mining operations, and manufacturing productivity, [2, 5, 6].

REGRESSION FORECASTING METHODOLOGY

Our regression forecasting methodology is designed to be applied to a historical data set of multiple inputs and outputs variables from a set of comparable units, [3, 4]. Additionally, one output variable is assumed to be the principal/critical variable that will be needed to be forecasted, e.g., sales, production, or demand.

Since the data set we studied has a relatively small number of inputs and outputs we adjust our procedure and eliminate the initial stepwise regression. As a result, the first step is to run a DEA for each comparable unit. We use these efficiency scores as surrogate measures of the unique emphasis of the variables and of performance. Using a principal/critical output variable as the regression dependent variable, all the input variables plus the DEA efficiency score as regression independent variables, we run the regression. This regression model with the DEA efficiency variable should be superior, i.e., should have a significantly lower standard error of the mean and increase R^2 , to the regression model without the DEA efficiency score variable.

EXAMPLE

Seiford and Zhu, [7], applied DEA to the 55 U.S. commercial banks that appeared in the *Fortune* 1000 list in April 1996. The DEA input variables were the number of employees, assets, and stockholder's equity; and the DEA output variables were revenue and profit. The selection of these variables were "based on *Fortune's* original choice of factors for performance characterization", [7].

We retrieved the same *Fortune* 1000 list of U.S. commercial banks from 2003 to 2006. We ran similar DEA models, i.e., same input and output variables as Seiford and Zhu, for 2003 to 2005. Table 1 below lists the frequency distribution of the DEA efficiency scores for these years. As shown in Table 1, these efficiency scores are rather dispersed.

Interval	YEAR		
	2003	2004	2005
100	2	4	6
90.01 to 99.99	3	1	2
80.01 to 90	6	5	9
70.01 to 80	6	10	6
60.01 to 70	9	4	3
50.01 to 60	2	2	2
40.01 to 50	1	2	
<40		1	

Table 1. Frequency distribution of the DEA efficiency scores for each year.

Using the DEA efficiency scores as an input and revenue as our primary output variable, we ran regression models for 2004 to 2006. The basic regression equation used was: $\text{Revenue}(t) = \text{Employees}(t-1) + \text{Assets}(t-1) + \text{Equity}(t-1) + \text{DEA}(t-1)$; where $t = 2004, 2005, 2006$; (we refer to this model as w/DEA). Additionally we ran the same regression without the DEA efficiency score variable (we refer to this model as NoDEA).

Tables 2 and 3 summarize the regression models results with R^2 values and standard errors. The w/DEA models were consistently better than the NoDEA models. In terms of R^2 values, the NoDEA models in each year had extremely high R^2 values. The w/DEA models only slightly increase the R^2 values; averaging only .1% improvement.

R^2	YEAR		
	2004	2005	2006
NoDEA	99.26	99.89	99.63
w/DEA	99.48	99.90	99.70
Difference	0.21	0.01	0.07

Table 2. The regression R^2 values for each year and for the two models.

Standard Error	YEAR		
	2004	2005	2006
NoDEA	2163.41	1042.98	2394.12
w/DEA	1862.84	1019.64	2191.75
Decrease	300.57	23.34	202.37
% decrease	13.89	2.24	8.45

Table 3. The regression standard errors for each year and for the two models.

The standard error values for the w/DEA models, in Table 3, had a more significant improvement; averaging 8.19% decrease in the standard errors.

Table 4 summarizes the residual results by displaying the maximum and minimum residual for each model. In each case, the w/DEA regression models performed better than the NoDEA regression models.

	2004			2005			2006		
	NoDEA	w/DEA	Improvement	NoDEA	w/DEA	Improvement	NoDEA	w/DEA	Improvement
Max	6346.03	3756.31	2589.72	1930.83	1699.98	565.30	8697.21	6902.34	1794.87
Min	-5469.5	-4638.14	-1906.04	-1787.71	-1720.04	-700.79	-3343.11	-3085.70	-1928.11
Average			102.38			6.28			33.74

Table 4. Residual analysis for each year and for the two models.

CONCLUSIONS

In this paper, we applied a new regression forecasting methodology to forecasting comparable units. This approach included in the regression analysis a surrogate measure of the unique weighting of the variables and of performance. This new variable is the relative efficiency of each comparable unit that is generated by DEA. The results of applying this new regression forecasting methodology including a DEA efficiency variable to a data set demonstrated that this may provide a promising rich approach to forecasting comparable units. We plan to perform further testing with other data sets, some with more comparable units and more years of data.

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] Klimberg, R.K. and D. Kern, "Understanding Data Envelopment Analysis (DEA)," *Boston University School of Management Working Paper*, pp. 92-44, 1992.
- [3] Klimberg, R. K., Lawrence, K. D., and S.M. Lawrence, "Improved Performance Evaluation of Comparable Units with Data Envelopment Analysis (DEA)," *Advances in Business and Management Forecasting*, Volume 5, Elsevier Ltd., pp. 65-75, 2008.

[4] Klimberg, R. K., Lawrence , K. D., and S.M. Lawrence, "Forecasting Sales of Comparable Units with Data Envelopment Analysis (DEA)," *Advances in Business and Management Forecasting*, Volume 4, JAI Press/North Holland , pp. 201-214, 2005.

[5] 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.

[6] Seiford, L. M. and R. M. Thrall, "Recent developments in DEA: the mathematical programming approach to frontier analysis." *Journal of Econometric*, 46, pp. 7-38, 1990.

[7] Seiford, L. M. and J. Zhu, "Profitability and Marketability of the Top 55 U.S. Commercial Banks," *Management Science*, Vol. 45, No. 9, Sept. 1999, pp. 1270-1288.

TIME SERIES ANALYSIS OF NORTHERN FUR SEAL FEEDING AND BEHAVIOR

James Bishop, Bryant University, 1150 Douglas Pike, Smithfield RI 02917

Email: jabishop@bryant.edu, Phone: 401-232-6356

Cheryl Clark, New England Aquarium, 74 Warren Rd., Plainville, MA 02037

Email: cclark@neaq.org, Phone: 508-695-8706

ABSTRACT

This paper analyzes the relationship between food intake, weight, and motivation for three northern fur seals in captivity at the New England Aquarium. The motivation for the study is for the trainers to better understand how the fur seals will react to changes in their dietary intake. There is more than one week lag time between a change in dietary intake and a significant weight gain or loss. In addition, motivation of the seals is negatively correlated with their weight.

Keywords: fur seal, feeding, behavior, weight.

INTRODUCTION

On May 23, 2006, the New England Aquarium welcomed three adult female northern fur seals on loan from the Mystic Aquarium in Connecticut [4]. Initially, six months of daily data was collected on one of these fur seals (Cordova) by the trainers at the New England Aquarium over the period July 1, 2006 through December 31, 2006. Analysis was performed relating the effect of food intake (independent variable) on weight and motivation (dependent variables).

We now have a full set of 2007 data for all three fur seals relating food intake and weight, and the results of this relationship are presented in this paper. Motivational figures for 2007 are not complete and so only the initial findings relating food intake to motivation are analyzed within. Data covering an entire year allows us to verify the initial findings, and study fur seals through seasonal cycles including the annual molting process.

INITIAL FINDINGS

Initially, six months of data was collected on one female adult fur seal. Feeding occurred once per day in the morning, measured in calories. The seals were weighed at the end of the day, measured in kilograms – weights were recorded every 2 – 5 days. Finally, motivational ratings were established on a 5-point scale (1 to 5) based on specific skill performance tests administered daily by the trainers. The Aquarium gives seal shows every day, and the response to tasks performed during the shows was included in the analysis.

Regression analysis was initially performed based on time-lag data between the time of the feeding and the subsequent weight gain and changes in motivation. The initial model was

$$W_i = \beta_0 + \beta_1 * L(k) + e_i; \quad (1)$$

where $L(k)$ is the average caloric intake per day over the most recent k days.

Minitab was used to perform simple linear regressions based on lag periods for k = 3,6,9,12,15,18. P-values for the β_1 coefficient are summarized in the following table:

K	P-value	K	P-value
3	.523	12	.013
6	.283	15	.032
9	.094	18	.085

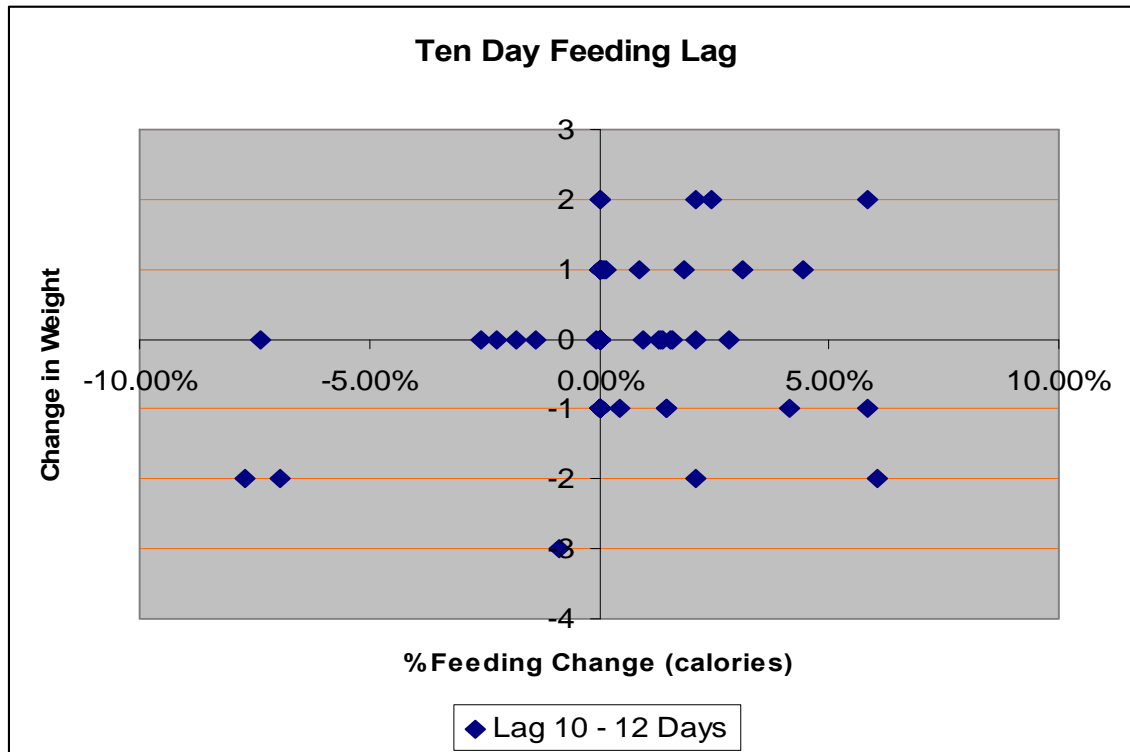
The findings suggest a surprising delay in the response variable based on feedings approximately 9 - 12 days prior. In order to get a closer look at the lag period, we change the regression model to group the feedings into three day lag periods. The adjusted model is

$$W_i = \beta_0 + \beta_1 * L_3(1) + \beta_2 * L_3(4) + \beta_3 * L_3(7) + \beta_4 * L_3(10) + \beta_5 * L_3(13) + e_i; \quad (2)$$

where $L_3(k)$ is the average caloric intake per day over the 3-day period with k as the most recent day.

Running a best subset regression, the significant lags can be determined with some care since there is significant collinearity present in the analysis. The coefficients change a bit as variables come in and out of the model. However, the $L_3(10)$ variable has a stable coefficient with a low p-value ($<.02$). Therefore, it is apparent that a full 12 days of feeding has a significant impact on the weight of this particular fur seal. A graphical comparison is given in Table 1 below.

Table 1



Additionally, a simple regression analysis was performed relating weight over this six month period (independent variable) to the daily motivation rating (dependent variable). This analysis uses data only on the days that Cordova was weighed. The results significantly showed a negative relationship. In other words, a hungry fur seal is a motivated fur seal.

The following is a brief summary of results based on the first 6-months of data analysis:

- The prior 12 days of feeding has a significant impact on weight gain
p-value = 0.01, $r^2 = 0.45$.
- The prior 9 days of feeding does not seem to have a significant impact on weight gain - p-value = 0.09.
- Weight on a given day has a clear negative relationship to motivation - p-value < 0.01.

FULL DATA MODEL

The full data model included data on three adult female fur seals at the New England Aquarium for the full calendar year 2007. The seals are all fully adult (ages 9, 10, and 20) and so there is no growing process to consider. Data was recorded on each of the three fur seals in a similar manner to the data collected in the initial study. Caloric intake was recorded daily and weights were taken every 2 – 5 days.

Weather and water temperature are not considered a factor in the study since water temperature is kept relatively constant between 55° and 62° Fahrenheit. Air temperature fluctuates more dramatically, but the seals are only out of the water approximately 10% of the time. We did not adjust for this in the study.

Each seal has a molting period near the end of the year (two of the months between September and December) in which they are less active and tend to gain weight. The trainers adjust caloric intake during these periods to offset the inactivity. Also, there is less activity during the winter months of January through March.

The regression analysis was conducted using an indicator variable to reflect periods of more limited activity. The indicator variable was set to one for molting periods, as well as, other winter months of limited activity. The use of only one indicator variable could be considered a limitation of the analysis since it assumes that the two types of inactivity are approximately the same.

The regression model used for the full data analysis is:

$$\Delta W_i = \beta_0 + \beta_1 * W_{i-1} + \beta_2 * \Delta W_{i-1} + \beta_3 * \Delta L_3(k) + I_m + e_i; \text{ where} \quad (3)$$

ΔW_i is the percentage change in weight from one weighing to the next,
 $\Delta L_3(k)$ is the 3-day average percentage change in caloric intake with day k being the most recent day in the averaging, and
 I_m is the indicator variable for periods of inactivity.

The prior weight and prior change in weight were considered for each of the seals' observed weights. This model stabilizes the regression analysis, preventing correlated residuals as can be shown by the Durbin-Watson statistic. The full results are summarized in Table 2.

The full data model has a time lag response that is not as long as what was observed in the initial study, except in the case of Chainsaw. For her, the regression using a lag of $\Delta L_3(10)$ has a lower p-value during periods of inactivity. Additionally, both Ursula and Cordova had low p-values for periods of inactivity using $\Delta L_3(10)$.

Table 2 – using $\Delta L_3(7)$ as the lag period

	β_0	β_1	β_2	β_3	r^2	p-value for β_3	D-W
Cordova	-3.20	0.046	-0.238	15.78	0.18	0.003	1.58
Chainsaw	-4.72	0.046	-0.182	12.85	0.22	0.067	1.96
Ursula	-12.06	0.166	-0.332	17.10	0.35	0.001	1.78

CONCLUSION

There is statistically significant evidence that weight gain in adult female fur seals is dependent upon food intake from 7 – 9 days prior. In addition, during periods of inactivity, this lag time may be up to 12 days. This information may be used to help the trainers control the diet of the fur seals, particularly because there is an effect on their motivation.

Only one fur seal was studied with regard to a relationship between motivation and weight. There was a statistically significant negative correlation. When the data for the other two fur seals (Ursula and Chainsaw) is made available, the correlation analysis should be performed to verify the consistency of this result. One practical application of this analysis is that the trainers can adjust caloric intake in order to influence weight gain and motivation level. This can help them understand how the fur seals will respond, particularly during shows.

One of the seals, Chainsaw, is significantly older than the other two and this may have contributed to a weaker result with regard to the lag time for weight gain. The correlations and p-values suggest that the lag period is may be closer to 12 days for her.

DISCUSSION

The New England Aquarium trainers have used the information presented in this paper to adjust feeding levels for the seals. Their feedback, based on the time lag presented here, is positive and they are getting the desired response in terms of desired weight and motivation levels.

This paper made some rough adjustments for periods of inactivity that could be improved upon. Also, a larger and more diverse sample would give us wider applicability for the results.

REFERENCES

- [1] COSEWIC, "Assessment and Update Status Report on the Northern Fur Seal (in Canada)". Canadian Wildlife Service. Ottawa, ON, 2006.
- [2] Dickey, David A. "Regression with Time Series Errors". Proceedings of 23rd Annual SAS Users Group International Conference, 1998.
- [3] Gentry, Roger L. "Behavior and Ecology of the Northern Fur Seal". Princeton University Press, 1997.
- [4] The New England Aquarium. Available at: <http://www.neaq.org/about/pressrelease/>, 2006.

STATISTICS IN THE WORKPLACE: THE PBA VS PISCATAWAY

Sheldon Epstein, Seton Hall University, South Orange, NJ 07079, 973 761- 9252, epsteish@shu.edu
Yonah Wilamowsky, Seton Hall University, South Orange, NJ 07079, 973 761- 9239, wilamoyo@shu.edu
Bernard Dickman, Hofstra University, Hempstead, NY 11549, 516 463- 5721, acsbhd@hofstra.edu

ABSTRACT

The State of New Jersey Public Relations Employment Commission recently rendered a decision in a dispute over the Piscataway Police department's procedure for promoting individuals to the rank of Sergeant. One important component of the case was how to properly interpret the results of a 1999 Sergeant's Promotion Exam. This paper gives a brief history of the promotional process and offers the data and statistical analysis submitted by both the Plaintiff (Patrolmen's Benevolent Association- PBA) and the Police Administration. The Case provides an excellent tutorial for beginners and practitioners on how to properly apply some elementary but powerful statistical concepts. **Key words: Case Study, Statistics, Correlation, Rank**

CASE BACKGROUND

In 1999 promotion to the position of Sergeant in the Piscataway Police Department was based on a written test, an oral interview and seniority. The written test was given first and was conducted by an outside agency chosen by the Police Administration. Only applicants passing the written exam with a grade of 70 or more (out of a possible 100 points) were permitted to continue on to the oral interview. The oral interview was conducted and scored in-house by the Piscataway Police Administration and each interviewee was assigned a grade between 0 and 100. The written and oral exam scores were then weighted at 50% each and added to the applicants number of years of seniority (up to a maximum of 10) to yield a final grade. The anticipated available Sergeant positions were to be filled from the applicants' list in descending order of final scores.

PLAINTIFF'S STATISTICAL CASE

Table 1 gives the actual results, in descending order of scores, from the 1999 Sergeant's Exam in which 8 positions were available and 23 applicants passed the written exam. The Piscataway PBA alleged that the oral test scores were manipulated by the Piscataway Police administration, knowing the results of the written exam in order to ensure a final ordering of candidates that was to the Administration's liking.

To demonstrate the "manipulation" of the numbers, the Plaintiff expanded Table 1 by adding three columns which rank all of the applicants who passed the written test from 1 to 23 in the individual areas of written scores, oral scores and seniority. Table 2 lists these ranking results. A brief filed by the PBA then asserted:

TABLE 1: 1999 SERGEANT’S TEST RESULTS

<u>Applicant</u>	<u>Written Score</u>	<u>Weight 0.50</u>	<u>Oral Interview</u>	<u>Weight 0.50</u>	<u>Seniority</u>	<u>Final Score</u>
1	87	43.5	90.2	45.1	10	98.60
2	86	43.0	85.0	42.5	10	95.50
3	88	44.0	81.6	40.8	10	94.80
4	72	36.0	91.2	45.6	10	91.60
5	79	39.5	84.2	42.1	10	91.60
6	87	43.5	77.2	38.6	7.75	89.85
7	78	39.0	80.4	40.2	10	89.20
8	80	40.0	75.0	37.5	10	87.50
9	90	45.0	64.0	32.0	10	87.00
10	80	40.0	70.0	35.0	10	85.00
11	82	41.0	68.0	34.0	10	85.00
12	84	42.0	65.0	32.5	9.75	84.25
13	78	39.0	68.6	34.3	10	83.30
14	84	42.0	60.8	30.4	10	82.40
15	72	36.0	79.2	39.6	6.5	82.10
16	74	37.0	68.0	34.0	10	81.00
17	87	43.5	53.2	26.6	9.75	79.85
18	81	40.5	57.4	28.7	10	79.20
19	83	41.5	55.2	27.6	10	79.10
20	86	43.0	51.8	25.9	10	78.90
21	74	37.0	65.2	32.6	6.5	76.10
22	72	36.0	58.4	29.2	10	75.20
23	74	37.0	53.0	26.5	10	73.50

“The rankings are separated into 2 groups of eight candidates and a group of 7 candidates. The first group of 8 candidates represents the initial individuals who were promoted. What is clear by simply “eyeballing” the results, is that there is a correlative relationship between the “Oral” ranking ... and the “Final” ranking... The “Oral” ranking which had a weighted value of 50% was within one to three points of the final promotional outcomes. Statistically revealing is that for the most part, the other groups also had a correlation between the oral ranking and the respective candidates final ranking.”

The Plaintiff pointed out that a similar relationship between final ranking and written ranking is absent and concluded:

“One can aver that how well a candidate scores on the “Oral” category is significantly more important in determining their “Final” rank. This phenomenon reeks of manipulation and impropriety given the fact that both components are weighted a 50%.”

TABLE 2: 1999 SERGEANT’S TEST RESULTS BY RANKING

<u>Candidate</u>	<u>Written Rank</u>	<u>Oral Rank</u>	<u>Seniority Rank</u>	<u>Final Rank</u>
1	3	2	1	1
2	6	3	1	2
3	2	5	1	3
4	21	1	1	4
5	15	4	1	5
6	3	8	21	6
7	16	6	1	7
8	13	9	1	8
9	1	16	1	9
10	13	10	1	10
11	11	13	1	11
12	8	15	19	12
13	16	11	1	13
14	8	17	1	14
15	21	7	22	15
16	18	12	1	16
17	3	21	19	17
18	12	19	1	18
19	10	20	1	19
20	6	23	1	20
21	18	14	22	21
22	21	18	1	22
23	18	22	1	23

To quantitatively buttress the “eyeballing” argument, the PBA offered Table 3 which lists Coefficients of Correlation (CC) for each of the different segments of data previously identified:

TABLE 3: CORRELATION COEFFICIENTS USING FINAL RANKINGS

<u>Candidates</u>	<u>With Respect To Oral Rank</u>	<u>With Respect to Written Rank</u>
1 through 8	0.79	0.47
9 through 16	-0.31	0.72
17 through 23	-0.13	0.81

Using these numbers the Plaintiff concluded.

- The CC for the ranking of the first 8 candidates with respect to the oral test and the final rankings is .79 while the CC for the ranking of the same 8 candidates with respect to the written test and the final ranking is a lower .47. This

demonstrates that although both the Oral and written exams were ostensibly given the same weight of .5, the Oral Exam had a considerably greater affect in determining who was being promoted to Sergeant, and

- The CC's for the ranking of the 2nd and 3rd groups of candidates with respect to the oral test and the final rankings are somewhat negative while the CC for the ranking of the same 2 groups of candidates with respect to the written test and the final ranking are highly positive.

DEFENDANT'S PERSPECTIVE

The Defendant rejected the PBA's methodology and inferences on the following grounds:

- In presenting its case the PBA offered no motivation for changing the results of the different examinations from actual test scores to rankings. In fact, changing from numerical scores to rankings loses significant information. The Correlation formula used to develop Table 3 is meant for actual numerical data, not for ranked data. Table 4 gives the corrected CC's based on the numerical data.

TABLE 4: CORRELATION COEFFICIENTS USING FINAL NUMERICAL VALUES

<u>Candidates</u>	<u>With Respect to Oral Score</u>	<u>With Respect to Written Score</u>
1 through 8	0.69	0.53
9 through 16	-0.29	0.79
17 through 23	-0.26	0.90

Note the .79 CC for oral ranking with respect to final ranking has been reduced to .69 and the .47 for written ranking has been increased to .53, so that the .32 difference in CC's has been cut in half to .16.

- The PBA offered no rationale for decomposing the 23 Applicants into 3 groups (8, 8 and 7 respectively) rather than 2 groups of 8 and 15 respectively (i.e. those promoted and those not promoted). In conducting a statistical analysis large groups of data are preferable to smaller ones. This case perfectly illustrates the point. Had the analysis been done with only 2 groups, the CC for Candidates 9 through 23, Table 4 becomes:

TABLE 5: CORRELATION COEFFICIENTS USING FINAL NUMERICAL VALUES

<u>Candidates</u>	<u>With Respect to Oral Score</u>	<u>With Respect To Written Score</u>
1 through 8	0.69	0.53
9 through 23	0.56	0.52

Note that the 2 negative values for the Oral tests are now over +.5 and the CC's for the written tests dropped sharply. In fact, the CC for Oral slightly exceeds that for Written for the applicants who were not promoted. Again, the volatility of CC with respect to the amount of data used (i.e. the switching of CC from negative to positive values as the number of items in the data set goes from 7 and 8 to 15) clearly illustrates why small subsets of data should not be used and that when all of the data is analyzed together the differences that allegedly point to "manipulation" disappear.

- The PBA's underlying assumption is that: It "reeks of discrimination" if it can be shown that the score on the Oral test is more important in determining which candidates finish at the top of the list is wrong. It can easily be shown that assuming the Written and Oral tests are both administered fairly and both scores are weighted at 50%, the more important factor in determining the people who finish at the top of the candidates list will be the Oral exam. The key to this analysis is the fact that only candidates receiving a grade of at least 70 on the Written exam are eligible for promotion. The same is not true for the Oral exam, i.e. candidates failing the Oral test (score under 70) are still eligible for promotion. Table 6 lists the Means, Standard Deviations and Ranges of each column of Table 1.

TABLE 6: MEANS, STANDARD DEVIATIONS, RANGES

	<u>Written</u>	<u>Oral</u>	<u>Seniority</u>	<u>Final Score</u>
Mean	80.8	69.7	9.6	84.8
Standard Deviation	5.7	12.2	1.1	6.8
Range	72 to 90	51.8 to 91.2	6.5 to 10	

For the Written exam the Mean for all of the candidates is 80.8, the SD is 5.7 and the grades range from a low of 72 to a high of 90 i.e. 18 points. While this set of numbers is not as compact as the Seniority numbers the distance from the Mean grade to the highest or lowest grade is only about 9 points. The Oral exam grades are the most dispersed of all with a Mean Grade of 69.7, a SD of 12.2 and a Range of grades from 51.8 to 91.2, i.e. 39.4 points, because:

- No grade, regardless of how low, was eliminated,
- The difference between the highest and lowest grades on the Oral exam is so much greater than the difference between the highest and lowest passing grades on the Written exam.

Anyone scoring at the upper end of the Oral exam is guaranteed to do considerably better than anyone scoring at the lower end of the Oral exam. For example, any candidate who outscores another by more than 18 points on the Oral exam is almost assured of placing ahead of the other candidate regardless of what grade either received on the Written exam (since the spread between the highest and lowest Written exam grade is only 18).

- The consistency of the data and the reasonableness of the selection of the 8 people who were actually promoted can be seen in Table 7. Table 7 differs from Table 1 in that all candidates who “failed” the Oral exam (i.e. received a grade of under 70) were deleted. Note that of all of the original applicants, only 10 passed both the written and oral tests and only 2 were not promoted: one had the lowest passing score on the Written exam and the other had the lowest passing score on the Oral exam.

TABLE 7: CANDIDATES WHO PASSED THE WRITTEN AND ORAL EXAMS

<u>Weight</u>	<u>100%</u>	<u>50%</u>	<u>50%</u>	<u>Final</u>
<u>Candidates</u>	<u>Seniority</u>	<u>Written Score</u>	<u>Oral Interview</u>	<u>Score</u>
1	10	87	90.2	98.60
2	10	86	85.0	95.50
3	10	88	81.6	94.80
4	10	72	91.2	91.60
5	10	79	84.2	91.60
6	7.75	87	77.2	89.85
7	10	78	80.4	89.20
8	10	80	75.0	87.50
10	10	80	70.0	85.00
15	6.5	72	79.2	82.10

Finally, recalculating the CC’s of Table 3 for all 10 candidates who passed both the Written and Oral exams yields Table 8:

TABLE 8: CORRELATION COEFFICIENTS USING FINAL NUMERICAL VALUES

<u>Candidates</u>	<u>With Respect to Oral Score</u>	<u>With Respect To Written Score</u>
10	0.71	0.65

Thus the initial major CC discrepancy reported by the PBA for the Written and Oral exams has been reduced from .28 (.79 to .47) to less than .06 (.71 to .65).

CONCLUSION

Based on these and other considerations the Arbitrator found in favor of the Defendant.

QUALITY SURVEY PROBLEMS

Robert J. Richardson, Iona College, 715 North Avenue, New Rochelle, NY 10801-1890,
(914) 637-7726, rrichardson@iona.edu

ABSTRACT

The annual initial quality survey of automobile buyers reports the number of problems found in the first 90 days of ownership. Since consumers use this information in purchasing an automobile, it is important that the information be clear and concise for the public to understand. In 2006, the surveys were changed to combine two measures of quality (number of defects/malfunctions and design) into a single number. In prior years, only a single measure, defects, was included. This paper addresses issues with the accuracy of aggregate values for the information in these surveys and the consistency over the years. The conclusion makes suggestions for improvements.

Keywords: Quality, Surveys, Statistics

INTRODUCTION

Each year, numerous automobile quality surveys are reported by several information services. Driven by the results of these studies, auto manufactures have come a long way in ensuring that vehicles are made better. The results of the J. D. Power's Initial Quality Surveys [7] of consumers and the awards for best car in a segment are used extensively in advertising and influence purchasing decisions. In 2006, the survey [6] was redesigned to capture problems experienced by owners in two distinct categories—quality of design and quality of production (defects and malfunctions). This paper addresses several issues with the accuracy in aggregating the quality data by model into nameplate values and then into country summaries for the information in these surveys and the consistency over the years. The conclusion recommends separate surveys to improve the information content and to make it more easily understood.

ACCURACY OF THE AGGREGATED VALUES

The surveys collect information by individual car model, the lowest level of the aggregation pyramid in Figure 1. In this example, the models G6, G5, Vibe, and others are combined into the nameplate, Pontiac. Then, the nameplates of Pontiac, Cadillac and others are aggregated into General Motors. Finally, manufacturers are aggregated into countries or areas of the world. The domestic manufacturers are the combined values for GM, Ford and Chrysler. At each level, the information is reported and comparisons are made. The model quality values are used to award/identify the three best (lowest number of problems) by car segment (subcompact, compact, midsize, etc.). The value by nameplate is presented in a table for review.

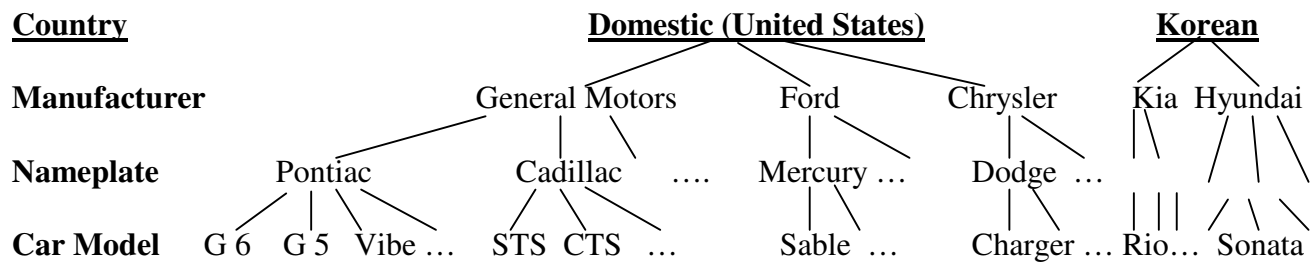


FIGURE 1. Levels of Aggregation Pyramid

To examine the accuracy of the process used in the surveys to combine these values, there are four methodologies that could be used in making these calculations for each country. These include: (1) Equally weighted by manufacturer, (2) Weighted by the number of cars sold by each manufacturer in the United States, (3) Weighted by the global sales of each manufacturer, and (4) Average of the survey values returned for each manufacturer

The 2004 survey [4] provides an opportunity to identify which method is used in calculating aggregate values. The easiest country to evaluate is Korea, since there are only two auto companies, Kia and Hyundai. For 2004, Korean cars had an average value is 117 PP 100 (Problems Per 100). From the 2004 report [4], KIA had 153 PP 100 and Hyundai had 102 PP100. Since the number of surveys received by company is unknown, it is not possible to compute the value of the fourth method. Also, since none of the Korean cars were made in the United States in 2004, there are no plants that could be included with the domestic cars value.

Figure 2 presents the Korean value calculated by using an equal weight for each company, by the US sales and by global sales. Since the awards are distributed and used in advertising in the United States, the sales should be based on U.S. sales. The 2003 sales were used in the calculation, because the report was issued on April 28, 2004. Since none of these methods yield the value of 117, the conclusion is that the values are calculated incorrectly based on the number of surveys returned to J D Power.

	<u>Values in Report[4]</u>	<u>Equally Weighted</u>	<u>Weighted by US Sales</u>	<u>Weighted by Global Sales</u>
KIA	153	50%	37% (237,471)	34% (858,697)
Hyundai	102	50%	63% (400,221)	66% (1,650,034)
Korean Value	117.0	127.5	120.9	119.4

FIGURE 2. Calculation of the Korean Average Value

Taking the average of all surveys misrepresents the value for Korea, since it fails to represent the distribution of the number of cars purchased by consumers. If the same methodology was applied to computing the company value from its models, a similar error in the values is generated. These values need to be calculated correctly using the latest sales data.

CONSISTENCY

By definition, consistency is the agreement with what has already been done or the conformity with previous activity. This is even more important when the information is distributed to the public. If the number of problems per vehicle varies widely from year to year without an explanation, then the sampling methodology may not be capturing the true values. Thus, consistency is one measure of the reliability of the results reported in Initial Quality Survey. First, the plant level values are examined, followed by the nameplate breakdowns and finally the awards by individual model.

Assembly Plant Awards

The top three plants in North and South America with the number of problems per 100 cars for last seven years are presented in Figure 3. The 2006 change in methodology does not affect the plant statistics since "Plant awards are based solely on defect counts" [7]. However, there is a significant decrease in the number of defects reported in 2006 and 2007. For example, the GM plant went from 85 to 42 or 49% decrease. These changes require an explanation.

	<u>2001</u>	<u>2002</u>	<u>2003</u>	<u>2004</u>	<u>2005</u>	<u>2006</u>	<u>2007</u>
1 st	Toyota 96	GM 93	GM 87	GM 74	GM 85	GM 43	Ford* 35
2 nd	Honda 101	GM 98	GM 88	GM 91	GM 89	Chrysler47	GM 42
3 rd	Toyota 101	GM 100	Ford 94	Ford 92	GM 90	Toyota47	Honda 44

* Indicates Platinum Award winner (Overall lowest number of defects for ALL plants)

FIGURE 3. Assembly Plant Quality Awards [1] [2] [3] [4] [5] [6] [7].

Nameplate Rankings

In 2006, the methodology changed with the introduction of the quality of design added to the quality of production measured in the number of problems reported by consumers. This modification was justified by recognizing that "New vehicles today are often packed with new technologies that unfortunately can be complicated and frustrating for the average consumer when their integration is not well executed. In the eyes of consumers, design flaws can have as much of an impact on their perceptions of quality as can a defect. Yet, many manufacturers have tended to address quality solely on the plant floor without considering design factors. Based on both design quality and production quality considerations, the study finds that automakers can vary widely in their performance on these two components" [6].

The number of problems from manufacturing operations is a quantitative value, a count of the number of defects, an integer number. The quality of design is a qualitative estimate based on the consumer's judgment of the cars design. Combining a qualitative feeling with an exact count is incorrect. These are two different measures that should not be combined.

The quality of design focuses on one dimension of design, the technology used in cars. The rating that a consumer gives to this measure is dependent upon many factors. If the respondent is young, new technology may be easier for them to deal with compared to older drivers. A bias is also created by the number options provided on the auto. Cheaper cars have substantially fewer options with less settings per option. The impact on BMW is used to illustrate the problem.

“Without considering both quality factors, one might fail to recognize vehicles that are, in fact, excellent in certain ways. For example, BMW vehicles have among the fewest defects and malfunctions, along with Toyota. But BMW approaches controls and displays in a way that creates some problems for customers, leading to more design-related problems overall than Toyota” [6]. The impact of adding the quality of design is illustrated in Figure 4 for BMW. For BMW, there is a 49% increase in 2006 compared to decreases 9% and 12% in the prior years. The increase comes when the BMW plants had a significant decline in 2006 and 2007. Is the increase a result of more or less problems in the factory, the design of the car, or both? These data are not consistent.

		<u>2003</u>	<u>2004</u>	<u>2005</u>		<u>2006</u>	<u>2007</u>
BMW	Ranking	8th	12th	3rd		27th	20th
	PP 100	118	108	95		142	133
	Yearly % Change		-9%	-12%		49%	-7%

FIGURE 4. Impact of Change in 2006 [3] [4] [5] [6] [7]

Further inconsistency is given by the impact of this change on the statistics by nameplate in Figure 5. Prior to the change, the average number of problems and the range decreased each year. After the addition of design, the mean increased to a value greater than the two prior years and the number of nameplates less than the average decreased. The mean value increases from 2006 to 2007. This is the first increase in the mean in eight years. These changes indicate that there is a shift in the distribution. The exact impact is difficult to evaluate, since the two quality measures are combined.

<u>Year</u>	<u>1998</u>	<u>1999</u>	<u>2000</u>	<u>2001</u>	<u>2002</u>	<u>2003</u>	<u>2004</u>	<u>2005</u>		<u>2006</u>	<u>2007</u>
Lowest						76	87	81		91	91
Highest						225	173	151		204	170
Range						149	86	70		113	79
Mean	176	167	154	147	133	133	119	118		124	125
Number of Nameplates Above Average						15	16	15		14	13

FIGURE 5. Values Over Years by Nameplate [1] [2] [3] [4] [5] [6] [7]

Initial Quality Surveys of customers ranked between 35 and 37 nameplates based on the number of problems per vehicles and quality of design after 2006. The quality information was segmented by global region (Domestic, European, Japanese, and Korean) by J. D. Power in 2004 [4]. Figures 6 counts the number and percent of brands by region for the top fifteen best nameplates, since there are usually fifteen that are better than average quality nameplates for each year in Figure 6. The average number of cars in each region is computed for the three years prior to the change and the two years after including design. The number one nameplate was Lexus for the first three years, with Porsche the leader for the last two years.

	<u>Domestic</u>	<u>Japanese</u>	<u>European</u>	<u>Korean</u>
2003	4 (27%)	5 (33%)	6 (40%)	0 (0%)
2004	4 (27%)	5 (33%)	5 (33%)	1 (7%)
2005	5 (33%)	5 (33%)	4 (27%)	1 (7%)
Average for 3 Years	4.3 (28%)	5 (33%)	5 (33%)	.7 (5%)
2006	7 (46%)	6 (40%)	1 (7%)	1 (7%)
2007	6 (40%)	5 (33%)	2 (14%)	2 (14%)
Average for 2 Years	6.5 (43%)	5.5 (37%)	1.5 (10%)	1.5 (10%)

FIGURE 6. Top 15 Nameplate Ranking Summary (J. D. Power IQS [5])

For the top fifteen, the gainers are the domestic autos with almost two cars and the Japanese with half a car on average after the change. These gains were at the expense of the Europeans who lost three and a half cars on average. Based on the top fifteen analyses, it appears that the Europeans have developed less desirable interfaces with technology by putting too much technology in their cars.

Nameplate Rankings

Finally, the individual models are ranked based on the awards by car segment (sub-compact, compact, midsize, etc.). The three best models are identified based on the least number of defects and design flaws. Figure 7 presents the count of the top cars in each category.

<u>Year</u>	<u># Awards</u>	<u>Domestic</u>		<u>Japanese</u>		<u>European</u>		<u>Korean</u>	
2001	16	5	31%	11	69%	0	0%	0	0%
2002	16	7	44%	9	56%	0	0%	0	0%
2003	16	8	50%	7	44%	1	6%	0	0%
2004	18	6	33%	10	56%	1	6%	1	6%
2005	18	7	39%	11	61%	0	0%	0	0%
5 Year Average		6.60	39%	9.60	57%	.40	2%	.20	1%
2006	19	5	25%	12	60%	1	5%	1	5%
2007	19	8	40%	6	30%	5	25%	1	5%
2 Year Average		6.50	33%	9.00	45%	3.00	15%	1.50	8%

Note: In 2006 and 2007, a single tie occurred increasing the total count by one to 20.

FIGURE 7. Number of First Place Autos in Segment

Contrary to the results in the nameplate analysis, the European and Korean manufacturers increased the number of those selected for an award, while the domestic and Japanese declined slightly. This is not unusual since the components of the nameplate averages many models that may have values that far exceed the mean of the nameplate.

Design has always been lacking in many cars when you consider another dimension, interior space. The tall person (over 6 feet 2 inches) has a problem in the front seat of many cars and an even bigger problem in the back seat. Design criteria needs to include comfort in terms of space as well as the design of the dashboard for readability, rear window for visibility, exterior for efficient airflow dynamics, etc. The selection of technology opens the door to all the other

aspects of design and should not be included with the number of defects to present an overall picture of initial quality.

CONCLUSIONS

Quality surveys have focused the consumer's attention on quality and forced the manufacturers to improve quality. The next step in improving the measurement of the initial quality of an automobile is to separate the number of defects/malfunctions and the quality of design into two surveys. The consumer needs both measures in making a purchase decision. By combining the two into one measure, the volatility of result is confusing and misleading.

Another survey of initial quality should focus on the number of recent recalls that a manufacturer has. In the past few years, the number of recalls on vehicles has increased to almost double the number of new cars sold. An overall measure of quality reflecting recall counts and percent of recalls by models sold should be included for the last five years. The recalls could be classified into different severity categories (fatal, dangerous, and minor). This would warn buyers of potential issues after purchase.

REFERENCES

- [1] Greywitt, Michael, and John Tews, "J. D. Power and Associates Reports: European Automakers nearly tie Asians on new-vehicle initial quality; domestic makers follow." *J. D. Power Press Release*, May 17, 2001, 1-9.
- [2] Greywitt, Michael, and John Tews, "J. D. Power and Associates Reports: Auto industry achieves 10 percent improvement in initial quality - largest one-year improvement sin 1997." *J. D. Power Press Release*, May 30, 2002, 1-9.
- [3] Greywitt, Michael, and John Tews, "J. D. Power and Associates Reports: New/redesigned models often record better initial quality than the vehicles they replace, but not in every case" *J. D. Power Press Release*, May 6, 2003, 1-9.
- [4] Greywitt, Michael, and John Tews, "J. D. Power and Associates Reports: Korean-branded vehicles overtake European and domestics in initial quality." *J. D. Power Press Release*, April 28, 2004, 1-8.
- [5] Greywitt, Michael, and John Tews, "J. D. Power and Associates Reports: Toyota motor corporation, General Motors Corporation garner most awards in 2005 initial quality study" *J. D. Power Press Release*, May 18, 2005, 1-7.
- [6] Tews, John and Peter Dadlani, "J. D. Power and Associates Reports: redesigned initial quality study shows that vehicle design plays a critical role in consumer perceptions of quality as defects and malfunctions" *J. D. Power Press Release*, June 7, 2006, 1-7.
- [7] Tews, John and Syvetril Perryman, "J. D. Power and Associates Reports: Ford motor captures most awards in 2007 initial quality study" *J. D. Power Press Release*, June 6, 2007, 1-7.

USING MIXED INTEGER NONLINEARLY CONSTRAINED OPTIMIZATION TO DO PENALIZED MAXIMUM LIKELIHOOD ESTIMATION FOR GARCH AND ARCH MODELS

Jingjing Lu, Drexel University, Department of Decision Sciences, 101 N 33rd St, Academic Building, Philadelphia, PA 19104, (215) 222-1848, jl466@drexel.edu

ABSTRACT

Compared to the traditional maximum likelihood regression approach, the penalized maximum likelihood estimation (PMLE) is a more rigorous method because of the adjustment for over fitting is directly built into the model development process, instead of relying on shrinkage afterwards. This paper illustrates the application of a nonlinear programming technique on PMLE to develop a prediction model for a Generalized Autoregressive conditional heteroskedasticity (GARCH) model based on Autoregressive Moving Average time series. Our numerical study demonstrates that the mixed integer nonlinearly constrained optimization method can accurately predict the parameters of time series data.

Key words: mixed integer optimization; nonlinear programming; penalized maximum likelihood estimation; Generalized Autoregressive conditional heteroskedasticity (GARCH) model; ARCH model

INTRODUCTION

In the time series literature, the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model plays an important role in volatility forecasting. It is a powerful tool used in finance to model asset price volatility over time. It is often observed that in long time series data on asset prices there are periods where variance is high and periods where variance is low. The GARCH econometric model introduced by Engle [3] shows that the variance of such a series is an autoregressive time series itself, often a linear one. This type of model considers the variance of the current error term to be a function of the variance of the previous time period's error term. GARCH relates the error variance to the square of a previous period's error. It is employed commonly in modeling financial time series that exhibit time-varying volatility.

The maximum likelihood estimation (MLE) is a popular statistical method used to estimate the model parameters and build statistical inferences time series models of a given data set. Actually, the maximum likelihood ratio theory contributes significantly towards parametric inferences. Engle [3], Pantula [4] and Bollerslev et al. [2] discussed the maximum likelihood (ML) based inference procedure for ARCH models under certain assumptions. Later Altay-Salih et al. [1] used constrained nonlinear programming and maximum likelihood techniques to build a GARCH volatility estimation model based on autoregressive time series. However, compared to the traditional maximum likelihood regression, the penalized maximum likelihood estimation (PMLE) is a more rigorous method, because the adjustment for over fitting is directly built into

the model development process, instead of relying on subsequent shrinkage. For most of the existing estimation methods, the lag parameters p and q are assumed to be known. Then the maximum likelihood or the least squares method is applied to estimate the model coefficients. But in real implementations, these values are usually unknown. So there appears to be a variable selection problem concerning the prediction of the values for p and q . In fact, this variable selection problem is much more involved than the estimation problem, especially in high-dimensional cases.

In this paper, we illustrate the application of mixed integer nonlinearly constrained programming on the PMLE method to develop a prediction model for the Generalized Autoregressive conditional heteroskedasticity model based on Autoregressive Moving Average time series, which is a more generalized model compared to [1]. This paper is arranged in the following way: at the beginning, we outline the mathematical deviation of the mixed integer nonlinear constrained optimization system of the penalized maximum likelihood problem. As a special case of the GARCH model, the Autoregressive conditional heteroskedasticity (ARCH) model's parameters can be easily estimated via our PMLE method. We also present the optimization results of our numerical example, as well as the comparison between the estimated parameters values and their actual values. Our results demonstrate that the penalized maximum likelihood estimation under the mixed integer nonlinearly constrained optimization framework can effectively predict the parameters of the time series data.

MATHEMATICAL MODEL

Suppose we observe $Y_i, i = 1, 2, \dots, T$ following a GARCH(m,n;p,q) model, which means that $Y = \{Y_t, t \in T\}$ is an ARMA (m,n) (autoregressive moving average) model with GARCH(p,q) errors:

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_m Y_{t-m} + \varepsilon_t - \phi_1 \varepsilon_{t-1} - \dots - \phi_n \varepsilon_{t-n} \quad (1)$$

where ε_t represents the random noise satisfying the following conditions:

$$E(\varepsilon_t | \varepsilon_{t-1}) = 0, \quad (2)$$

$$E(\varepsilon_t^2 | \varepsilon_{t-1}) \equiv \sigma_t^2 = c + \sum_{i=1}^p a_i \varepsilon_{t-i}^2 + \sum_{j=1}^q b_j \sigma_{t-j}^2. \quad (3)$$

Suppose the process is stationary and assume that Y_t and ε_t are normally distributed. Then for any t ,

$$E(Y_t) = \mu, \quad (4)$$

$$Var(Y_t) = \sigma^2, \quad (5)$$

$$Cov(Y_t, Y_{t-s}) = \gamma_s. \quad (6)$$

The values of μ, σ^2, γ_s do not depend on t .

Note that to ensure the second-order stationary condition, we need:

$$\sum_{i=1}^p a_i + \sum_{j=1}^q b_j < 1 \quad (7)$$

Given this assumption, asymptotically, the process will achieve the stationary condition. In fact, we can look at it in the following way: the process has been running for a long time before we start observing data, so that all the asymptotical phenomena are in effect.

Suppose r and u are the upper bounds for the parameters p and q , respectively. To estimate the parameters $\phi_i, i = 1, 2, \dots, m$, $\varphi_i, i = 1, \dots, n$, $a_i, i = 1, 2, \dots, r$, and $b_j, j = 1, 2, \dots, u$, together with p and q , we can use the penalized maximum likelihood method, which can be modeled as the following nonlinear optimization problem:

$$\text{Max} - \sum_{t=1}^T \log(\sigma_t^2) - \sum_{t=1}^T \frac{\varepsilon_t^2}{\sigma_t^2} - 2(p + q) \quad (8)$$

Subject to:

$$\left\{ \begin{array}{l} c + \sum_{i=1}^r a_i \left(\frac{|i-p+0.5| + i-p+0.5}{2|i-p+0.5|} \right) \varepsilon_{t-i}^2 + \sum_{j=1}^u b_j \left(\frac{|j-q+0.5| + j-q+0.5}{2|j-q+0.5|} \right) \sigma_{t-j}^2 = \sigma_t^2, \\ \sum_{i=1}^m \phi_i Y_{t-i} + \varepsilon_t - \sum_{j=1}^n \varphi_j \varepsilon_{t-j} = Y_t, \forall t = 1, 2, \dots, T \\ \sum_{i=1}^r a_i \left(\frac{|i-p+0.5| + i-p+0.5}{2|i-p+0.5|} \right) + \sum_{j=1}^u b_j \left(\frac{|j-q+0.5| + j-q+0.5}{2|j-q+0.5|} \right) \leq 1; \\ \sigma_t^2 \geq 0; \forall t = 1, 2, \dots, T; \\ c \geq 0; \\ a_i \geq 0, b_j \geq 0; i = 1, 2, \dots, r; j = 1, 2, \dots, u \\ r \geq p \geq 1; \\ u \geq q \geq 1. \end{array} \right. \quad (9)$$

In the above model, $-\frac{1}{2} \sum_{t=1}^T \log(\sigma_t^2) - \frac{1}{2} \sum_{t=1}^T \frac{\varepsilon_t^2}{\sigma_t^2}$ is the ordinary log-likelihood. For any fixed p and q , one can choose proper coefficients to maximize it and the maximized value is called the maximum likelihood. It can be seen that when p or q is increasing, the maximum likelihood is also increasing. To avoid the over-fitting problem, we add a penalty term $-(p + q)$ according to the Akaike information criterion. This assigns large penalties for large $p + q$ values. We can also use other penalties, such as $-(p + q) \log(T)/2$, which is the Bayesian information criterion or $-\log(p + q)(p + q)$ instead. Consequently we can still reformulate the problem as a nonlinear optimization problem.

Although the above model is not necessarily convex, it can be proved that the global maximum always exists. The proof is as follows: we know that for any t , when σ_t^2 or ε_t^2 (or p) goes to infinity, the log-likelihood always goes to negative infinity. Since the function is continuous, the global maximum exists. This is a standard result in the maximum likelihood methodology.

As a special case, our approach can easily be applied to an ARCH(m;p) model. In this case, the observation $Y = \{Y_t, t \in T\}$ satisfies

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_m Y_{t-m} + \varepsilon_t. \quad (10)$$

Now the random noises satisfies

$$E(\varepsilon_t | \varepsilon_{t-1}) = 0, \quad (11)$$

$$E(\varepsilon_t^2 | \varepsilon_{t-1}) \equiv \sigma_t^2 = c + \sum_{i=1}^p a_i \varepsilon_{t-i}^2. \quad (12)$$

Similarly, when the time series is stationary, the estimation of the parameters can be obtained via the following nonlinear optimization problem:

$$\text{Max} - \sum_{t=1}^T \log(\sigma_t^2) - \sum_{t=1}^T \frac{\varepsilon_t^2}{\sigma_t^2} - 2p \quad (13)$$

Subject to:

$$\left\{ \begin{array}{l} c + \sum_{i=1}^r a_i \left(\frac{|i-p+0.5|+i-p+0.5}{2|i-p+0.5|} \right) \varepsilon_{t-i}^2 = \sigma_t^2, \forall t = 1, 2, \dots, T; \\ \sum_{i=1}^m \phi_i Y_{t-i} + \varepsilon_i = Y_t, \forall t = 1, 2, \dots, T \\ \sum_{i=1}^r a_i \left(\frac{|i-p+0.5|+i-p+0.5}{2|i-p+0.5|} \right) \leq 1; \\ \sigma_t^2 \geq 0; \forall t = 1, 2, \dots, T; \\ c \geq 0; \\ a_i \geq 0; i = 1, 2, \dots, r; \\ r \geq p \\ p \geq 1. \end{array} \right. \quad (14)$$

NUMERICAL STUDY

In this part, we study the performance of our suggested method. First, a sequence of ARCH(2;2) time series data was generated by S-plus. The first two terms of the sequence were two independent standard normal random variables. We used the last 1000 data points as our observations. We dropped the first 14000 instances to guarantee the stationarity of the sequence. The theoretical values of this sequence's parameters were set to $m = 2, p = 2, \phi = (0.6, 0.2), a = (0.3, 0.1)$, and $c = 0.1$.

The following two figures show the time series data. Figure 1 shows all the data (from observation 1 to 15000) and Figure 2 only shows the data that we used (from observation 14001 to 15000).

Figure 1

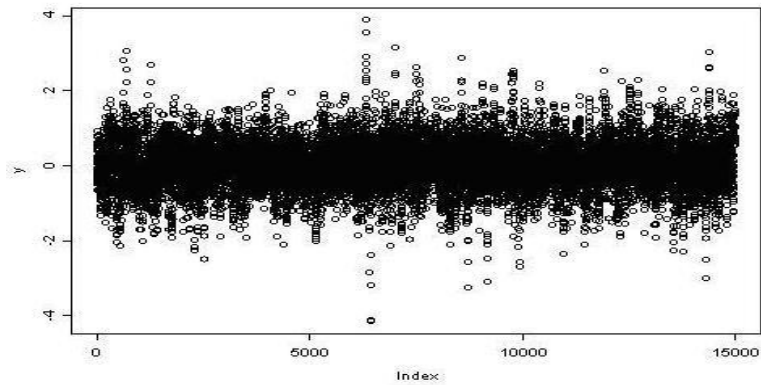
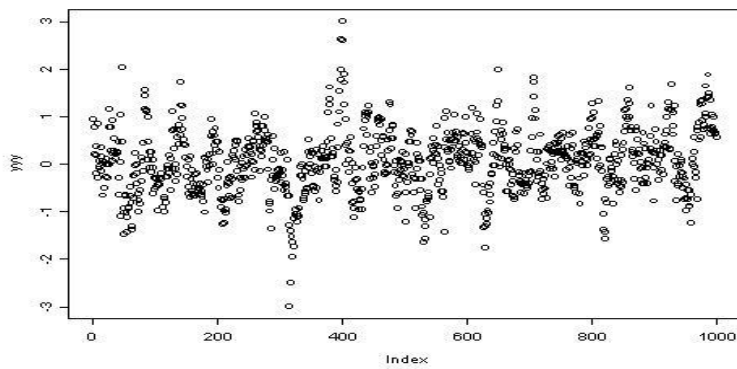


Figure 2



We used the AMPL language and the NEOS Server to solve the non-linear optimization problem, in order to estimate the parameters in question. The results are summarized in Table 1 (In reality, most of the ARCH models are with parameters p less than or equal to 4. Here we show the different objective function and parameter values of p from observation 1 to 9):

Table 1

P	1.0000	2.0000	3.0000	4.0000	5.0000	6.0000	7.0000	8.0000	9.0000
A1	0.3696	0.3458	0.3458	0.3458	0.3454	0.3454	0.3454	0.3454	0.3454
A2	0.0000	0.1176	0.1176	0.1176	0.1179	0.1179	0.1179	0.1179	0.1179
A3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
A4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
A5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
A6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
A7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
A8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
A9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
phi1	0.6158	0.6099	0.6099	0.6099	0.6083	0.6083	0.6083	0.6083	0.6083
phi2	0.2028	0.2047	0.2047	0.2047	0.2066	0.2066	0.2066	0.2066	0.2066
c	0.1072	0.0927	0.0927	0.0927	0.0911	0.0911	0.0911	0.0911	0.0911
object	867.3850	875.6400	873.6400	871.6400	869.9100	867.9100	865.9100	863.9100	861.9100

From the above results, we can see that the parameters are estimated quite precisely. When p equals 2, we obtained the maximum objective function value and the estimated parameters are very close to their actual values. Thus, using the nonlinear constrained optimization method, we accurately estimated the value of p . This means that the AIC rule is a proper choice of penalty in our scenario.

CONCLUSIONS

In the time series literature, the GARCH model is usually estimated by the maximum likelihood technique, based on the assumption that p , q are known. But in reality, the value of p and q are unknown. Thus the variable selection problem will become much more involved than the maximum likelihood estimation problem. In this paper, we solved this problem by applying a mixed integer nonlinearly constrained optimization approach to the penalized maximum likelihood estimation method. Our results demonstrate that nonlinear programming techniques can be a powerful tool for accurately predicting the relevant parameters of time series.

REFERENCES

- [1] Altay-Salih, Pinar, A., M.C., Leyffer, S.. "Constrained Nonlinear Programming for Volatility Estimation with GARCH Models". *SIAM Review*, 2003, 45, 485-503.
- [2] Bollerslev, T., Chou, R. Y., Kroner, K. F. "ARCH modeling in finance". *Journal of Econometrics*, 1992, 52, 5-59.
- [3] Engle, Robert F.. "Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation". *Econometrica*, 1982, 50, 987-1008.
- [4] Pantula, Sastry G.. "Modeling the persistence of conditional variances: A comment". *Econometric Reviews*, 1986, 5, 71-74.

AN ALTERNATIVE METHOD TO TEST THE RESIDUALS IN A REGRESSION MODEL

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@cob.rit.edu

ABSTRACT

In this paper we propose an alternative method to the Durbin-Watson (DW) test for the fitness of a regression model (see, for example, Makridakis [5, pp. 267-268; 303-304; 630-631] and Wilson & Keating [8, pp. 182-184; 234-236] for DW statistic). The proposed method tests whether the residual terms (i.e., $\text{actual}_t - \text{model}_t$) display any sign of non-randomness by comparing two variance estimators of the residuals. This test is a transformation of the DW statistic into a standardized normal statistic, $N(0,1)$, which is readily interpreted; and unlike the DW test does not require a special table. A numerical example is provided.

Keywords: Regression models, residuals, Durbin-Watson statistics, test for significance.

DISCUSSION

The Durbin-Watson (DW) statistic is a test statistic used to check for the presence of autocorrelation in the error terms in regression models. To test the significance of the DW statistic one must resort to DW tables. In this paper we are proposing an alternative test of significance to the DW test which is a transformation of DW statistic into a $N(0,1)$.

The proposed approach uses the significance test of two variance estimators of the residuals: one is the regular variance estimate, the other is the one computed using mean square successive differences (MSSD) – see , for example, Neumann, et al. [6], Hald [2, pp.357-360], Holmes and Mergen [3], [4]. The MSSD is defined as

$$\text{MSSD} = \frac{1}{(n-1)} \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2 \quad (1)$$

where n is the number of observations.

Using these differences, an unbiased estimate for the process variance is given by Hald [2] as

$$q^2 = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2 \quad (2)$$

Then the MSSD standard deviation is determined by taking the square root of q^2 . Roes, et al. [7] suggested a minor correction factor in estimating the standard deviation when the MSSD approach is used. This factor disappears as the sample size gets bigger. The significance of the difference between the conventional and MSSD variance estimates can be tested using the test given in Dixon and Massey [1, pp. 353-354].

$$z = \frac{1 - \frac{q^2}{s^2}}{\sqrt{\frac{n-2}{(n-1)(n+1)}}} \quad (3)$$

where the usual variance estimator (s^2) is

$$s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} \quad (4)$$

Values of z between ± 3 indicate that the difference between the two estimates is not significant, i.e., the data set seems to be random. Z values bigger than $+3$ and less than -3 indicate that the two estimates are significantly different and thus the data is not random. Z values bigger than $+3$ imply trend and long-term cycles in the data and values less than -3 indicate short term cycles. Thus we propose the use of this z statistic to check the randomness of the error terms in regression models. Within the context of the regression, X_i values are replaced by error terms (e_t), i.e., residuals. To calculate the proposed z statistic, two variance estimators will be calculated for the error terms, one using the MSSD approach and the other one using the regular approach as given in equation 4. Thus z values within ± 3 imply that the error terms do not show any sign of non-random patterns. Since z is $N(0,1)$ when n is bigger than 20, the use of z values between ± 3 gives about a 99.7% critical region for the test.

One can show the relation between the proposed test and the DW test as follows. DW statistics is given as

$$DW = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2} \quad (5)$$

where e_t is the residual associated with the observation at time t . The DW statistic lies between 0 and 4, and its value around 2 implies no autocorrelation among the error terms. DW values near 0 indicate positive correlation, whereas values near 4 imply negative correlation. The DW statistic is very similar to the term $\frac{q^2}{s^2}$ given in equation 3 except the division by 2 in q^2 . We can re-write equation 3 as follows:

$$z = \frac{1 - \frac{DW}{2}}{\sqrt{\frac{n-2}{(n-1)(n+1)}}} \quad (6)$$

Multiplying both sides by the term in the numerator in equation 6 will yield

$$\sqrt{\frac{n-2}{(n-1)(n+1)}} (z) = 1 - \frac{DW}{2} \quad (7)$$

And multiplying both sides by 2 will result in equation 8 as follows:

$$(2) \sqrt{\frac{n-2}{(n-1)(n+1)}} (z) = 2 - DW \quad (8)$$

Thus we can express DW as a function of z which is given below in equation 9.

$$DW = 2 - 2 \sqrt{\frac{n-2}{(n-1)(n+1)}} (z) \quad (9)$$

Equation 9 clearly shows the relation between the DW statistics and the proposed z statistic. As can be seen in the equation, when z is at its desired value of 0 for randomness of the error terms, DW will be equal to its desired value of 2. This approach does not intend to replace the DW statistic test nor does it claim to be a better approach. It is proposed just as an alternative test to check the randomness of the error terms in regression and autoregressive models. Its advantage, as it is stated in the beginning, is that it does not require a special table to check the significance of the z values. We believe this will be a practical method to quickly check the randomness of the error terms in regression models. However, the drawback of the proposed approach is that it does not indicate the sign of the correlation between the error terms, if there is significant correlation.

EXAMPLE

The example data set (see the Appendix) is from a machine shop producing break drums. The data, which is disguised to maintain confidentiality, is on the diameter of sequentially machined hubs. The test for stability indicated that the time series data given in the Appendix is not stable. This was tied to the fact that there is an automatic control mechanism in the shop which adjusts the process automatically. This, in return, creates an autocorrelation in the process. The autocorrelation function given in Figure 1 supports this.

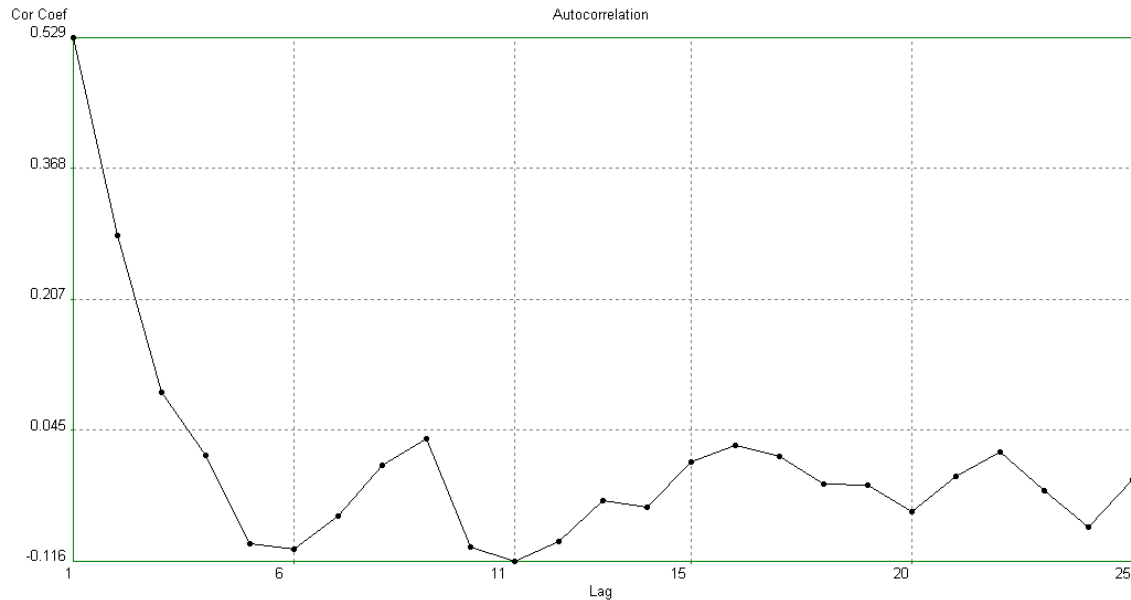


Figure 1. Autocorrelation chart on hubs diameter.

Thus we fit a first order autoregressive (AR) model to model the autocorrelation in the data (equation 10) and calculated the error terms, e_t , (equation 11):

$$\hat{X}_t = 1.395 + 0.53X_{t-1} \quad (10)$$

$$e_t = X_t - \hat{X}_t \quad (11)$$

where X_t is the actual observation and \hat{X}_t is the estimate at time t , respectively.

To test the significance of this AR model we used both the DW test and the proposed method on the error terms (i.e., the actual – the model given in equation 10 for each time period). The DW statistic and the z value from the proposed method using equation 6 are given below:

DW = 2.003 and the $z = -0.077$.

Both the DW statistic and the z value from the proposed method indicate that the error terms are random, which implies that the first order AR model is a good fit for this data.

CONCLUSION

In this paper we proposed an alternative method to the Durbin-Watson (DW) test for the fitness of a regression model. The proposed method is a practical method to quickly check the randomness of the error terms in regression models, since it does not require a special table to check the significance of the z values computed by the equation 6.

APPENDIX

Data on hubs diameter

2.966	2.966	2.974	2.970	2.960	2.958
2.966	2.960	2.988	2.959	2.950	2.977
2.964	2.966	2.981	2.953	2.964	2.982
2.953	2.967	2.979	2.940	2.968	2.977
2.953	2.963	2.981	2.959	2.971	2.981
2.958	2.983	2.980	2.980	2.983	2.976
2.965	2.986	2.972	2.981	2.968	2.986
2.973	2.960	2.944	2.976	2.986	2.972
2.978	2.982	2.958	2.969	2.981	2.981
2.987	2.979	2.958	2.961	2.998	_____
2.977	2.978	2.952	2.953	2.974	
2.964	2.976	2.975	2.951	2.982	
2.966	2.976	2.976	2.948	2.987	
2.976	2.952	2.970	2.971	2.981	
2.967	2.973	2.968	2.960	2.976	
2.981	2.967	2.955	2.979	2.958	
2.983	2.980	2.944	2.972	2.950	
2.989	2.973	2.944	2.979	2.960	
2.979	2.970	2.973	2.976	2.964	
2.966	2.980	2.963	2.971	2.968	
2.965	2.983	2.960	2.976	2.978	
2.961	2.968	2.956	2.965	2.980	
2.968	2.955	2.965	2.967	2.984	
2.976	2.942	2.961	2.960	2.978	
2.973	2.955	2.968	2.962	2.966	
2.967	2.960	2.972	2.963	2.972	
2.964	2.962	2.973	2.961	2.975	
2.976	2.974	2.964	2.953	2.974	
2.965	2.968	2.966	2.964	2.977	
2.968	2.963	2.957	2.968	2.979	
2.968	2.967	2.968	2.959	2.978	
2.987	2.966	2.985	2.960	2.973	
2.976	2.963	2.976	2.972	2.971	
2.963	2.967	2.988	2.968	2.962	
2.964	2.964	2.984	2.990	2.956	
2.990	2.966	2.981	2.979	2.973	
2.995	2.964	2.967	2.964	2.968	
2.975	2.966	2.969	2.967	2.972	
2.980	2.987	2.978	2.978	2.969	
2.980	2.991	2.975	2.952	2.986	
2.966	2.975	2.964	2.962	2.977	
2.965	2.983	2.958	2.953	2.965	

REFERENCES

- [1] Dixon, W.J. and Massey, F.J., *Introduction to statistical analysis*, New York, NY: McGraw-Hill, 1969.
- [2] Hald, A. *Statistical theory with engineering applications*, New York, NY: Wiley, 1952.
- [3] Holmes, D.S. and Mergen, A.E., "An alternative method to test for randomness of a process," *Quality and Reliability Engineering International*, 1995, Vol.11, 171-174.
- [4] Holmes, D.S. and Mergen, A.E. "Estimating potential capability of an unstable process," *The 2004 Northeast Decision Sciences Institute Meeting Proceedings*, Atlantic City, NJ, March 24-26, 2004, 290-292.
- [5] Makridakis, S., Wheelwright, S.C. and Hyndman, R.J. *Forecasting: methods and applications*, 3rd edition, Danvers, MA: Wiley, 1998.
- [6] Neumann, J.V., Kent, R.H., Bellinson, H.R. and Hart, B.I., "The mean square successive difference," *Annals of Mathematical Statistics*, 1941, Vol.12, 153-162.
- [7] Roes, K.C.B., Does, R.M.M. and Schuring, Y., "Shewhart-type control charts for individual observations," *Journal of Quality Technology*, 1993, Vol.25, No.3, 188-198.
- [8] Wilson, J.H., Keating, B. *Business forecasting*, 5th edition, Boston, MA: McGraw-Hill/Irwin, 2007.

THE FORGOTTEN POWER OF A TEST

Alan Olinsky, Bryant University, (401) 232-6266, aolinsky@bryant.edu
Phyllis Schumacher, Bryant University, (401) 232-6328, pschumac@bryant.edu
John Quinn, Bryant University, (401) 232-6392, jquinn@bryant.edu

ABSTRACT

In this paper, we discuss the importance of power considerations in statistical hypothesis testing. Statistical power analysis determines the ability of a study to detect a meaningful effect size, where the effect size is the difference between the hypothesized value of the population parameter under the null hypothesis and the true value when the null hypothesis turns out to be false. Specifically, we focus on a common statistical hypothesis testing example, a test of hypothesis of means concerning one sample. For this case, we examine the power of the test at varying levels of significance, sample sizes, standard deviations and effect sizes, using several statistical software packages. These power analyses are also available for other cases such as means and proportions of two samples, analysis of variance and regression analysis. An illustration of power in regression is included.

INTRODUCTION

In the statistical application of hypothesis testing, the significance level is always emphasized but often the power of the test is ignored. The significance level is the probability of rejecting a true null hypothesis, that is, committing a Type I error. Hypothesis tests are often called significance tests and the practitioner is encouraged to reject the null hypothesis if the evidence is overwhelming. In the absence of statistical significance rather than accepting that the null hypothesis is true, it is commonplace to conclude that it cannot be rejected thus avoiding committing a type II error, which is the error of accepting a false null hypothesis. Indeed the common analogy of comparing hypothesis testing to the U.S. legal system which assumes that the accused is innocent beyond a reasonable doubt or acquitted in the absence of enough evidence, encourages the consideration of the level of significance to the exclusion of considerations of the alternative hypothesis being true.

It is convenient to limit testing to consideration of levels of type I error rather than deal with power or increases in sample size; however, with large samples it is easy to reach statistical significance. This is one reason why traditional hypothesis testing does not occur in data mining where sample sizes are so large that everything is statistically significant. A statistical difference is not the same as a practical difference. With small samples, practical differences may not be statistically significant and vice-versa, with very large samples statistical differences may not be of practical importance. Several authors in diverse fields have addressed the importance of these considerations. Thomas and Juanes [7] indicate that in biology, for a large enough sample size “any statistical hypothesis test is likely to be statistically significant, almost regardless of the biological importance of the results.” Even though it is more elusive, it is important not to ignore the possible alternative hypotheses and effect sizes (differences between the hypothesized

and actual value of a parameter) and the calculation of the power of a test, that is, the probability of rejecting a false null hypothesis.

Using appropriate statistical software packages and relevant examples, we illustrate the computation of power for hypothesis testing. Specifically, we use Visual Statistics 2.0 [2] for demonstration purposes and SAS to compare and contrast the ease of calculating power and/or sample size effects [6].

Many texts give guidelines as to how many cases are needed when performing hypothesis test but these are very broad. Our analysis will be more specific and will generate sample sizes necessary under different scenarios to achieve an acceptable power level, which is often recommended to be a minimum of 0.8 [1]. As pointed out by Lenth [3], it is important to note that this type of analysis should be carried out prospectively rather than retrospectively. As he points out, “using retrospective power for making an inference is a convoluted path to follow. The main source of confusion is that it tends to be used to add interpretation to a non-significant statistical test.”

Our analysis will focus on a common statistical hypothesis testing example, that is, a test of hypothesis of a population mean. We will examine the power of the test at varying levels of significance, sample size, effect size, and standard deviations. We present our results in tabular and graphical format.

DEMONSTRATION

In the following example and illustrations from Visual Statistics 2.0 [2], power curves are presented which illustrate how the power is affected by changes in the sample size, levels of significance and at different levels of the effect size.

The voltage of a new battery is supposed to be 1.5 volts. The population standard deviation is believed to be 0.1 volts. A random sample of 36 new batteries is taken and the voltage of each battery is measured. What is the power of a two-tailed test at the .01 level of significance if the true mean is 1.45 volts?

Figure 1 represents a null hypothesis of $\mu=1.5$, with a two tailed test at a level of significance of 0.01 and a standard deviation of 0.1. In the middle curve, with $n=36$, when the alternative is 1.45, the power is 0.66. The top curve, with an increase in sample size to 72, shows that at the same alternative the power increases to almost one and with a decrease in sample size to 18, illustrated in the bottom curve, the power decreases to around 0.3. One can also observe that if the effect size is zero, the power is 0.05 because the null hypothesis is true and as the effect size increases, in both directions, the power increases.

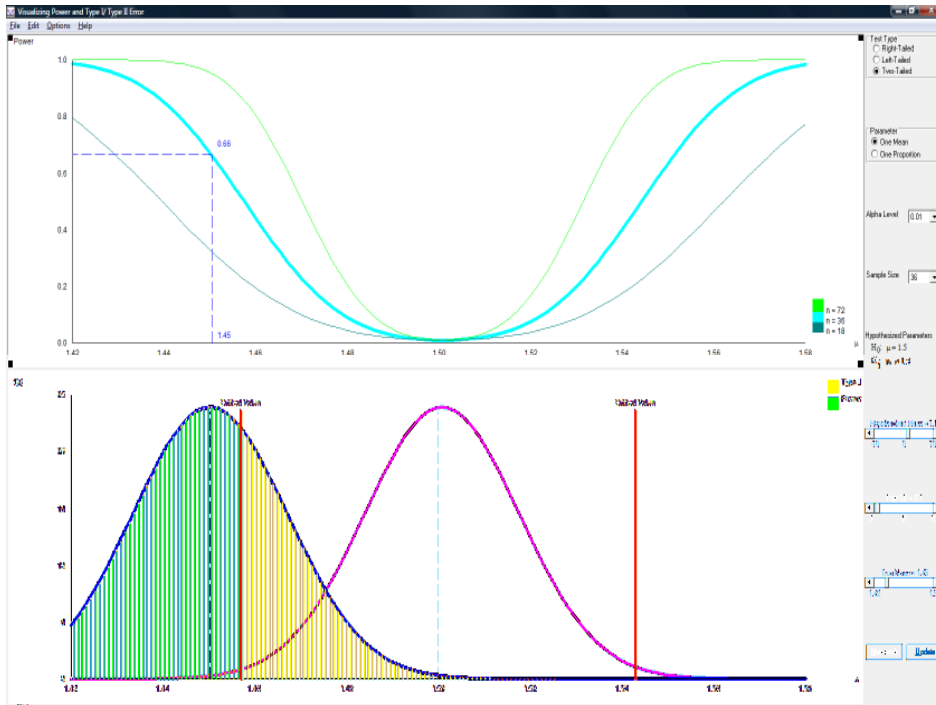


Figure 1: Family of Power Curves Varying Sample Size

In figure 2, we look at levels of significance of 0.01, 0.05, and 0.1 and keep n at 36 and the standard deviation at 0.1. We notice here that as the level of significance increases the power increases at a specific effect size, with again larger power at larger effect sizes.

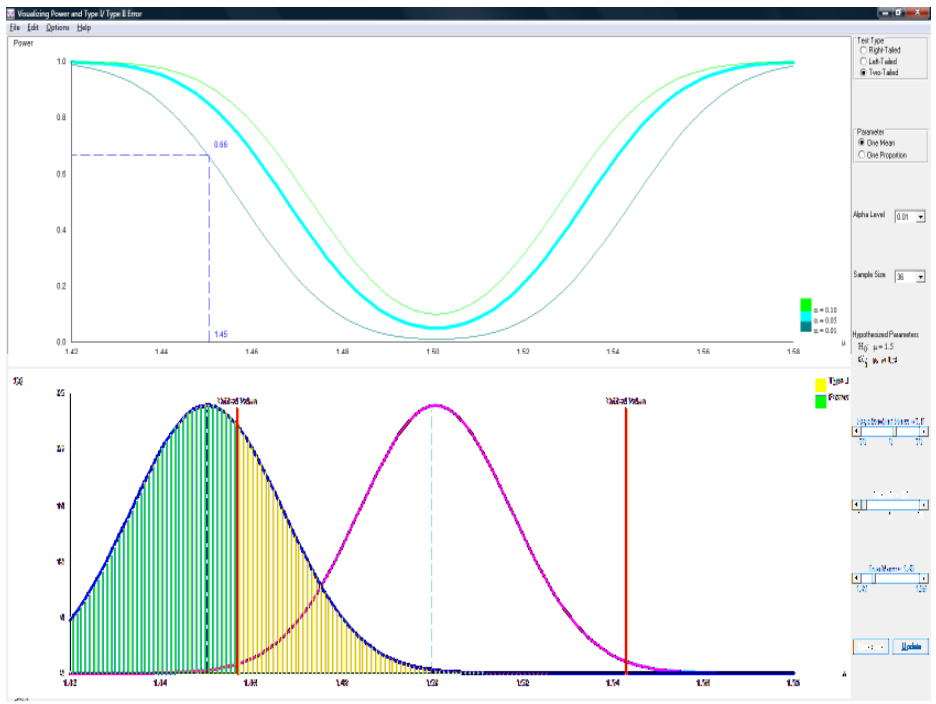


Figure 2: Family of Power Curves Varying Level of Significance

Figure 3 illustrates how the power changes with changes in the standard deviation. Here, the level of significance of 0.01 and the sample size of 36 are kept constant and the power curves are calculated for standard deviations of 0.05, 0.1 and 0.2. In this case for a given effect size, as the standard deviation increases, the power decreases.

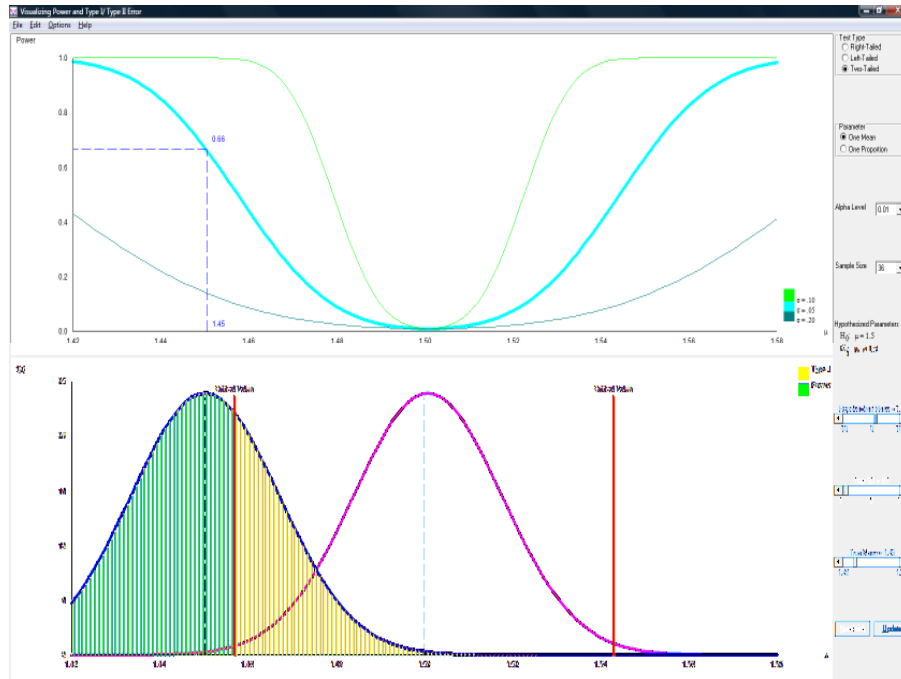


Figure 3: Family of Power Curves Varying Standard Deviation

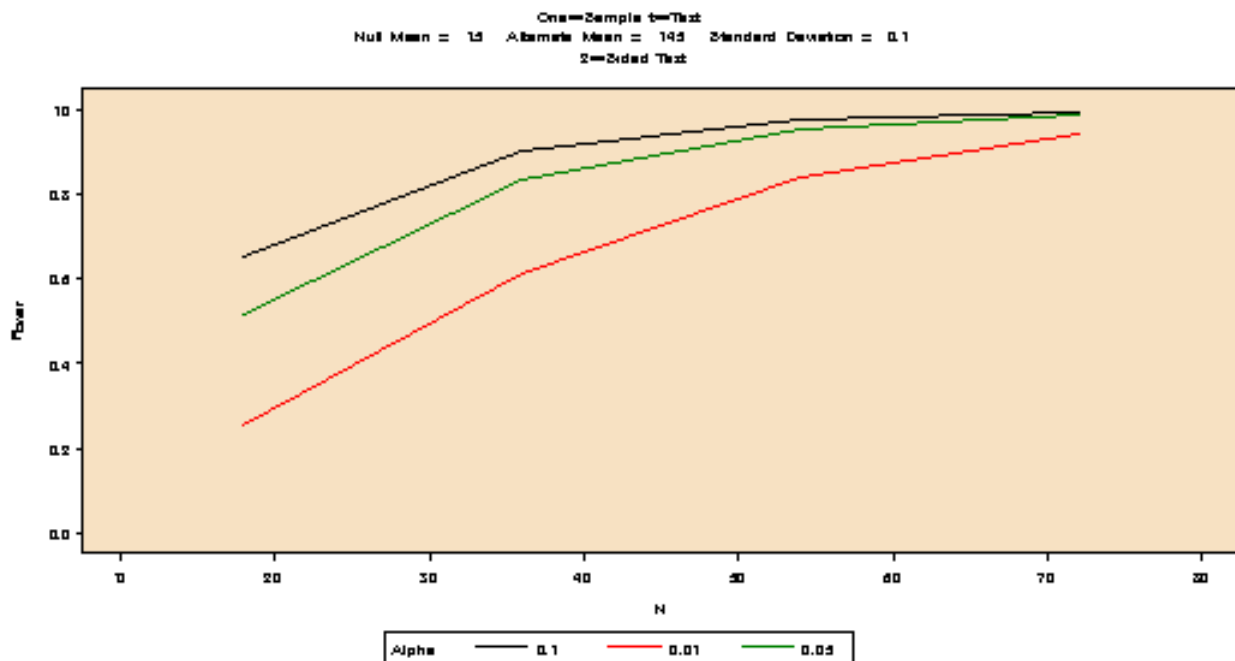


Figure 4: SAS Graph of Power Varying Sample Size and Level of Significance

Using the same example, in figure 4, SAS provides output for power curves at different levels of significance (0.01, 0.05, and 0.1) and for various sample sizes from 18 to 72. It is also interesting to note that in order to reach the power of 0.8, with a significance level of 0.01, approximately 54 observations are needed, at a significance level of 0.05, only 36 observations are needed and at a significance level of 0.1, fewer than 30 observations. MINITAB can also easily be used to obtain these calculations and graphs [4].

The above illustrations provide an example of the most common hypothesis testing procedure. Power analysis can be computed for more advanced testing procedures. There are several interesting examples of advanced power calculations and additional software package power procedures in the article, ‘Understanding the Statistical Power of a Test,’ by Hun Park [5]. There are even free calculator applications available on the Web to perform power analysis with different models including regression. For example, in figure 5, we have included the output from one such website, www.danielsoper.com. In this example, we entered a level of significance of 0.05, the number of independent variables of 3, an R-sq of 0.13 and a power of 0.8. The resulting required sample size was found to be 76.



Figure 5: Output for Sample Size Calculation for Multiple Regression

CONCLUSION

Statistical power analysis allows one to determine the ability of a study to detect a meaningful effect size. According to Hair [1], a researcher should strive for a power level of at least 0.8 for a study design. Hair also notes, as we have shown in the above demonstrations, more stringent significance levels (e.g., .01 instead of .05) require larger samples to achieve the desired power level and smaller effect sizes always require larger sample sizes to achieve the desired power. In addition, we found that smaller standard deviations also increase the power of a test procedure.

Our goal in this paper was to emphasize for researchers the importance and ease of determining appropriate sample sizes when planning their research studies. Such proper planning will reduce the risk of conducting a study that doesn't produce useful results.

REFERENCES

- [1] Hair, J. et. al. *Multivariate Data Analysis*. Upper Saddle River, NJ: Prentice Hall, 2006.
- [2] Doane, D., Tracy, R., and Mathieson, K. *Visual Statistics 2.0*. Columbus, OH, McGraw-Hill Inc., 2001.
- [3] Lenth, R. "Some Practical Guidelines for Effective Sample Size Determination," *The American Statistician*, 2001, Vol.(55), 187-193.
- [4] MINITAB 15, Minitab Inc., State College, PA, 2007.
- [5] Park, H. "Understanding the Statistical Power of a Test," Retrieved from <http://www.indiana.edu/~statmath/stat/all/power/power.pdf> on January 16, 2008.
- [6] SAS Institute. *SAS/STAT User's Guide, Version 9.1*. Cary, NC: SAS Institute, 2004.
- [7] Thomas, L., Juanes, F., "The Importance of Statistical Power Analysis: An Example from *Animal Behaviour*," *Animal Behaviour*, 1996, Vol(52), 856-859.

ILLUSTRATION OF QUANTIFYING THE SUPPLYCHAIN PROCESS

Seetharama L. Narasimhan
College of Business Administration
University of Rhode Island, Kingston RI 02881 USA
narasimhan@uri.edu (401) 874-4190

ABSTRACT

Many operations and supply chain management texts treat the subject both in qualitative and/or quantitative perspective. Our objective is to show how the entire process can be quantified in every stage of the manufacturing. Our contribution to the research is that we were able to separate the backorder case from lost sales case in inventory analysis using IF function in EXCEL. We were also able to use Max function in Production Activity Control example for starting the next activity to succeed either the previous activity or until the activity in the previous job is completed. In addition, we were able to use the management science techniques to solve transportation problems.

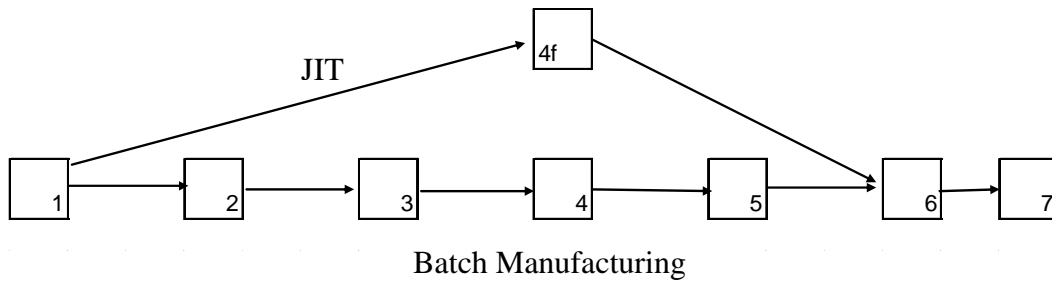
Subjects: Integration of processes, minimizing the total cost of products, effective manufacturing process, streamlining the manufacturing system

INTRODUCTION

Many operations and supply chain management texts treat the subject both in qualitative and/or quantitative perspective. Books listed in the Reference and many other books do a very fine job. Our objective is to show how this process can be quantified in every stage of the manufacturing. To start with, we divide the SCM process into: Continuous (JIT) manufacturing and Batch manufacturing. In either case, we need forecasting (medium and short term). While using continuous/JIT manufacturing, we can use the forecast for computing the number of Kanbans. In batch manufacturing, these forecasts provide inputs to capacity planning in medium term/aggregate planning which feeds into short term scheduling. Therefore, we start the manuscript with forecasting analysis followed by capacity planning, and production activity control. For both continuous/JIT and batch manufacturing we need to address the quality and process capability analysis and therefore we do so next.

DESCRIPTION OF THE PROCESS

We start with the Forecasting analysis and then we show how we can utilize the Inventory Management techniques with Quantity Discount models to choose vendors and follow subsequent analysis as indicated in the sequence exhibited below.



1. Forecasting
2. Sourcing
3. Transport to manufacturing facility
4. Capacity Planning
- 4f. JIT/Kanbans
5. Production Activity Control
6. Process Capability
7. Transport to warehouses/Retailers

1. Forecasting -Click on SCM Forecasting Analysis

2. Sourcing/Purchasing Decisions

Sourcing decisions can be made either using inventory management techniques or using assignment algorithm. Here we choose to utilize inventory management techniques. We will utilize the assignment algorithm while solving transportation problems. Both methods are based on cost.

Table 2 -Click on Inventory Sourcing Problem [2]

Two vendors are competing for your business. Which one will you choose?

H=\$2

S=20

Supplier A		Supplier B	
Quantity	Unit Price	Quantity	Unit Price
0-99	15	0-150	14.5
100-199	14	150-300	13.5
200 +	13.5	300+	13

It is a classic quantity discount problem. Find the best quantity & cost for each.

Table 3 Transport Components/Parts/Assemblies to Manufacturing Facility

-Click on SCM Transport to Facility Cost [4] [5]

Table 4a Capacity Planning [2] [3] - Click on SCM Chase Strategy

Aggregate planning provides capacity requirements which can also be translated to number of Kanbans in assembly line/continuous/lean manufacturing operations.

Chase Strategy			Regular	Regular	Hiring	Lay-off
Month	Plan	Actual	Hire	Layoff	Cost	cost
	1600				20	50
	1600				0	0
Jan	1400	1600	0	200	0	10000
Feb	1600	1400	200	0	4000	0
Mar	1800	1600	200	0	4000	0
Apr	1800	1800	0	0	0	0
May	2200	1800	400	0	8000	0
June	2200	2200	0	0	0	0
July	1800	2200	0	400	0	20000
Aug	1400	1800	0	400	0	20000
			Costs =		16000	50000
Total Cost =					\$66000	

We observe there is a change either in Hiring or Lay-offs when there a change in actual capacity occurs. Also observe the capacity changes frequently. Therefore, we need to reduce the number of Hires and Lay-offs as we show the results in Table 4d. Actually, we use mathematical programming techniques and minimize the numbers to one. Also when we use JIT it is also important that we keep the changes in capacity to a minimum.

Table 4b Capacity Planning- Click on SCM SOAP3

Actual Capacity = Average Capacity

If function was used to separate Inventory and backorder from FI

Negative Final Inv =			Backorder Case			
A	B	C	D	E	F	G
			Total Plan Cost			
Inventory Cost		20	64500			
Back Order Cost		50	16250			
		Sum	80750			

Month	Plan	Actual	B. I.	F. I.	20	100
Jan	1400	1775	200	575	575	0
Feb	1600	1775	575	750	750	0
Mar	1800	1775	750	725	725	0
Apr	1800	1775	725	700	700	0
May	2200	1775	700	275	275	0
June	2200	1775	275	-150	0	150
July	1800	1775	-150	-175	0	175
Aug	1400	1775	-175	200	200	0
Average	1775					
Sum Qty					3225	325
Costs					64500	16250
Total Cost					\$80750	

In Cell F12 we use =IF(E12>=0,E12,0). In cell G12 we use =IF(E12<0, 0, -E12) and copy to August in both columns

Table 4c Capacity Planning

Actual Production Capacity = Average Capacity

Negative Final Inv = Zero Begin Inv Lost Sales Total Cost

		Total Plan Cost	
InventoryCost	20	68000	
Lost Sales Cost	100	17500	
	Sum	85500	

Month	Plan	Actual	B. I.	F. I.	B.I	20	100
Jan	1400	1775	200	575	200	575	0
Feb	1600	1775	575	750	575	750	0
Mar	1800	1775	750	725	750	725	0
Apr	1800	1775	725	700	725	700	0
May	2200	1775	700	275	700	275	0
June	2200	1775	275	-150	275	0	150
July	1800	1775	-150	-25	0	0	25
Aug	1400	1775	-25	375	0	375	0
Total Cost =						3400	175
						\$85500	

Same as in the previous example but we add in cell FI Feb = IF(D40<0,-D40)

BI negative inventory becomes zero and proceed further as usual

And copy to July. Given: Aug Inventory is >= 0.

It is generally known that no reputable manufacturer will hire and lay-off workers frequently and spoils their reputation in the community. We need to reduce the number of hires and lay-offs more efficiently. In Table 4d we show how to minimize the number of hires and layoffs and yet don't sacrifice a lot.

Table 4d Capacity Planning- Click on SCM Hire&Fire

SCM_Hire&Fire H & R OPT 1

B. I.	200		
F. I.	200		
Hiring Cost	50	20000	
Layoff Cost	100	0	
Holding Cost	20	52000	
	Sum	\$72000	

Month	Plan	Average	Production	Hire	Layoff	FI
			1600	0.00	0.00	200.00
Jan	1400	1775	1600	0.00	0.00	400.00
Feb	1600	1775	1600	0.00	0.00	400.00
March	1800	1775	1600	0.00	0.00	200.00
April	1800	1775	2000	400.00	0.00	400.00
May	2200	1775	2000	0.00	0.00	200.00
June	2200	1775	2000	0.00	0.00	0.00
July	1800	1775	2000	0.00	0.00	200.00
August	1400	1775	2000	0.00	0.00	800.00
				400	0	2600.00

This is accomplished by using following equation:

$$I_t = I_{t-1} + \text{Production}_t - \text{Planned sales}_t$$

With constraints Inventory ≥ 0 , Hires and Lay-offs ≥ 0

To start: Leave Production, Hires and Lay-off columns as zeros and solve the optimization model. The solution provides actual production, hires and lay-offs and inventories. We find there was only one time hire and zero lay-offs. Even though the cost is \$72000 compared to Chase Strategy cost of 66000, we would think the firm would like the results provided by the optimization model as there are no constant hires and lay-offs.

Table 4e Multi Product/ Multi-Source Problem [4]

There is enough information given in the problem to describe the data.

Click on SCM SOAP5

Table 4f Computing # OF Kanbans- Click on Number of Kanbans

$$k = \frac{12.5 * 5 * (1 + .2)}{25} = 3$$

Using 5th month's demand of 2200 units, we compute the number of Kanbans required

Table 5 Production Activity Control [1] - Click on SCM Production Activity

We now proceed from planning to execution of the system.

Consider The following four jobs waiting to be processed

Job	Estimated Days			TaskTime		
	Step 1	Step 2	Step 3	Days	DueDate	CR
AB	3	4	3	10	22	0.454545
CD	5	2	2	9	21	0.428571
EF	2	2	4	8	8	1
GH	6	5	3	14	15	0.933333

We consider FCFS, EDD and CR rules Results and choose the sequence which provides the minimum lateness to complete the project as we are interested to serve customer's interest.

EDD: Earliest Due Date

FCFS: First Come First Served

CR: Critical Ratio Ratio of Task time to Due Date

Job	Step 1		Step 2		Step 3		Due Date	Late Days
	Start	End	Start	End	Start	End		
EF	0	2	2	4	4	8	8	0
GH	2	8	8	13	13	16	14	2
CD	8	13	13	15	15	17	21	0
AB	13	16	16	20	20	23	15	8

Average Lateness = **2.5**

As explained, we use **Max** function to determine the starting date of next job

FCFS	Step 1		Step 2		Step 3		Due date	Late Days
	Start	End	Start	End	Start	End		
AB	0	3	3	7	7	10	22	0
CD	3	8	8	10	10	12	21	0
EF	8	10	10	12	12	16	8	8
GH	10	16	16	21	21	24	15	9

Average Lateness = **4.25**

CR	Step 1		Step 2		Step 3		Due date	Late Days
	Start	End	Start	End	Start	End		
EF	0	2	2	4	4	8	8	0
GH	2	8	8	13	13	16	15	1
AB	9	11	13	17	17	20	22	0
CD	11	16	16	18	18	20	21	0

Average Lateness = **0.25**

We find the average lateness is smallest by using CR sequence.

Therefore, if customer waiting time is important, we should choose CR rule.

Max function chooses the later of the previous stage and previous job finishes.

6. SCM Quality/Process Capability Analysis [3]

– Click on Quality & Process Capability

7. Transportation Costs- from Warehouses and/or to Retailer Destinations

The problem was solved similar to the problem Transport to Factory in Item #3.

- Click on Monthly cost of Transport from Warehouse to Retailers

CONCLUSION

We show that we are able to quantify the entire SCM process.

REFERENCES

[1] Bozarth and Handfield (2005) *Introduction to Operations and Supply chain Management*. Prentice-Hall, NJ.

[2] Heizer and Render (2003) *Operations Management*, 7th ed. Prentice-Hall, NJ

[3] Chase, Jacobs and Aquilano (2006) *Operations Management*, Irwin, NY

[4] Narasimhan, McLeavey and Billington (1995) *Production Planning and Inventory Control*, 2nd ed. Prentice-Hall. NJ

[5] Ragsdale (2007) *Spreadsheet Modeling & Decision Analysis*, Thompson, Mason, Ohio

DO SUPPLY CHAINS EXIST IN VIRTUAL WORLDS?

Peggy D. Lee, Ph.D.
Pennsylvania State University – Great Valley
School of Graduate Professional Studies
Management Division
30 E. Swedesford Road
Malvern, PA 19355
PDL3@psu.edu
610-648-3289

ABSTRACT

Virtual worlds such as Second Life, Utopia Universe and Active Worlds are the next frontier in communication, social networking, and electronic commerce. Developed by their creators as multiplayer online role playing games (MMORPG), these 3D virtual worlds have grown exponentially, creating many opportunities and challenges for businesses, social scientists, educators, policy makers, and the legal profession. Virtual entrepreneurs have become real life millionaires selling everything from real estate to shoes. This paper presents preliminary results of an exploratory study into the question of whether the presence of virtual economies creates opportunities and challenges for supply chain design.

Keywords: virtual words, supply chain, Second Life, virtual business

INTRODUCTION

In the last year, virtual worlds have exploded into the landscape of business and education in a way that exceeds the explosive influence of the Internet in the early 1980s. Educational institutions from Harvard and MIT to for-profit entities have a presence in some virtual world, most of them in the Linden Labs-developed Second Life. Second Life is a 3D virtual world, created exclusively by its residents. The convergence of almost every media into one platform has been primarily responsible for this phenomenon. The business use of Second Life is dominated by collaborative activities such as meetings of virtual teams, employee training, customer support and sales activities [1].

Although real life businesses use Second Life for a multitude of purposes, the Second Life is estimated to have a \$30 million economy. As of October 2007, over \$1 billion have been invested in 35 virtual companies (<http://VirtualWorldsWeekly.com>, October 4, 2007) by firms such as Redpoint Ventures, Charles River Ventures, Intel and Rustic Canyon Partners. Two acquisitions totaled \$810 million of these investments: Walt Disney's \$700 million acquisition of Club Penguin and Intel's acquisition of 3D virtual worlds graphics technology company Havok.

Clearly, Second Life is no longer "just a game". The following sections describe Second Life and its economy, present the few scholarly works regarding virtual businesses, and provide the results of a pilot study of supply chain design in this virtual environment. The concluding

sections provide a research agenda for future work in the role of virtual worlds in supply chain design and development.

WHAT IS SECOND LIFE?

Second Life, launched in 2003 by Linden Labs (<http://www.lindenlabs.com>) of San Francisco, California, is an online, virtual 3D multi-user environment that functions as a virtual world. It is an immersive social interaction environment. It is not a game in that the actions of its residents are not predetermined by a set of rules. All of the content of Second Life is created by its residents. As of January 28, 2007, there were 12.2 million residents of Second Life, with over 50,000 of them being in-world business owners with positive cash flow (http://www.secondlife.com/whatis/economy_stats.php). Residents fall into three categories in Second Life: 1) those who build community; 2) those who create content; and 3) those who own and operate businesses. Social interaction is one of the primary reasons why people create avatars (a real person's virtual world persona) is to interact with others. Research suggests that avatars are extensions of their real counterparts and that people tend to treat their avatars as if they are in real social space. Avatars can be customized and changed over time through the use of different clothing, hair styles, shapes and skins, all of which can be purchased with Linden dollars (approximately 270 lindens per US dollar).

By and large, however, content creation is fabric of Second Life. Everything that exists is there because a resident or a group of residents created it. Collaboration is encouraged and promoted by the fact that creation often occurs in the open where anyone can make a suggestion. Residents spend a lot of time creating content in Second Life, estimated to be as much as 4 million hours a month.

The third category of residents, those who operate virtual businesses, is the subject of this paper. Because residents own everything they create, they can market and sell those creations to other residents. In fact, our interviews of resident entrepreneurs indicated that what began as a desire to create became a business as an afterthought. The most successful businesses in Second Life are real estate, social sites such as clubs, and fashion retailing. Service businesses include wedding and event planners and managers and psychologists.

REVIEW OF RELEVANT LITERATURE

The scholarly literature is non-existent with regards to the virtual business activities and ventures. This technology is so new that only anecdotal evidence exists as to its importance as an economy with its own currency exchange (Lindex), codes of conduct, and culture. The overwhelming majority of literature on virtual worlds is about the presence of real world entities in Second Life and other virtual worlds ([1], [2], [3], [4], [5]). Businesses such as IBM, Cisco Systems, Toyota, Vodafone, Pontiac, Dell, and many others are using Second Life to prototype new products; train employees, and conduct market research. For example, Toyota uses Second Life to conduct market research on its Scion brand (http://slurl.com/secondlife/Scion%20City/47/38/23/?title=Scion_City). You can design a new computer at the Dell Second Life site and have it shipped to your real home. Banks are using Second Life to interview prospective employers and prototype bank layouts. Not-for-profit firms

are also operating in Second Life. The Alzheimer's Society of Ontario has developed an exhibit with information about Alzheimer's. The Centers for Disease Control (CDC) uses Second Life to increase community awareness of health issues. Educational institutions are using Second Life as extensions of their distance and residence education efforts.

Nood & Attema (2006) published one of the first studies on Second Life, finding that 16% of residents surveyed considered themselves entrepreneurs, meaning that they claimed to transact business in Second Life and exchange Linden dollars for real currency. There is scant empirical evidence of how virtual businesses operate, especially from a supply chain perspective. This paper seeks to address this issue.

RESEARCH DESIGN

The nature of Second Life makes survey research problematic in that sampling involves approaching avatars to ask for their time to complete a survey. Many firms offer incentives to get people to come to a web site to complete a survey [6]. Consequently, I elected to complete an ethnographic approach [7] for data collection and analysis. One key research question was the focus of the data collection: how are virtual supply chains designed and maintained. This research question was answered by interviewing business owners about their operations. The researcher assumed the role of participant-observer, spending many hours in Second Life. Participation included attending Second Life events, interacting with other Second Life residents, learning to create content, becoming a landowner, operating an in-world business, and learning the Linden scripting language. Data included field notes, interview transcripts, in-world snapshots, and other documents such as listservs and blog postings. The researcher lived in Second Life as a resident for approximately 4 months prior to conducting the study.

The actual research interviews were conducted as part of a class project on virtual worlds in the authors graduate level operations management class. Student/avatar interviewers interviewed virtual entrepreneurs to learn about how business is conducted in a virtual world and compare it with how that same business type is conducted in a real world environment. In the process, the interviewers learned valuable information about virtual supply chains.

Between September and December 2007, 58 virtual entrepreneurs were interviewed about their business operations within Second Life. They were not asked their real names and they were assured complete anonymity. They were asked about how they procured the products they sold and how the facilitating goods were obtained to support their service businesses. They were also asked about other operational issues such as inventory policies and procedures, customer service, managing capacity, and hiring and scheduling of employees. The data was analyzed recursively in the sense that as new ideas emerged, additional questions were injected into the interviews to yield a richer data set.

Of the 58 business owners, the majority were in retail. The other business types were tours (1), dance clubs (1), concert venues (2), gaming arcades (1), hotels (1), auto dealerships (1), ice hockey leagues (1), and meeting planners (1).

PRELIMINARY FINDINGS

This section provides preliminary findings the interviews conducted. Although virtual supply chains exist in Second Life, they are not managed in the same way that they are in real businesses. Just as the Internet collapsed the layers of communication between and among organizations and their employees, Second Life has collapsed the supply chain from an echelon-oriented structure to a loose network of relationships between entrepreneurs. When asked how they managed their supply chains, most of the business owners we spoke to did not know what we meant and the terms had to be explained to them. This is because the majority of the SL business owners we interviewed were there to create content. Marketing and selling what they created was not their primary motivation. Many of them had come to Second Life purely to interact with others. Once there, they learned how to create content using the building tools and the Linden scripting language. Many of them made friends with other residents who taught them about life in a virtual world.

A typical retail business owner in Second Life is also the creator of the items he or she sells. Thus, the supply chain for the retail establishment is significantly shorter than it is in the real world. Supplier selection is simply done by using the Search mechanism of Second Life to find a source for a particular item. The business owner visits the shop of the supplier, buys the item and installs it in his or her establishment. Often the two business owners never interact with each other. In a highly immersive and interactive world, we found this phenomenon to be analogous to the fact that in the real world, the manufacturer will rarely meet the end consumer.

The retail business owner might procure a shop to sell his or her items, but in many cases the building skills required to make dresses for example, are the same to build a building. So, many retailers build their own shops rather than renting them from others. They do, however, procure (purchase or rent) land from landowners. Equipment used in the retail business also includes vendors, a device used to dispense the products that are sold. These are purchased from other retailers who specialize in making vendors for others to use.

Another unique aspect of business in Second Life is its robust service sector. Tours, event planners, wedding planners, counselors, psychologists, and entertainment are the most prevalent service businesses in second life. While they have supply chains, they too are shortened in that the facilitating goods used in the service business are procured from someone they may know. As described above, however, this is not likely to be the case. A wedding planner, for example, might develop a list of preferred suppliers for flowers, venues, wedding dresses, tuxedos, hair, shoes, etc. in the same way that a real wedding planner does. One of the more interesting interviewees was a woman who runs a tour business in Second Life. Her supply chain is by far the most extensive in that wherever she takes new residents on tour, she must gain access to the property and get permission to land her vehicles there. Landowners in Second Life can restrict who can put their objects on their land. This business owner had to arrange all of this in advance in the same way that a tour guide would need to in real life.

Implications for Supply Chain Design

Real life supply chains can be cumbersome and extensive, involving multiple layers and creating a network of organizations devoted to providing products and services to end user customers [8]. Virtual supply chains, however, are lean and compressed with only one or two layers. What does this mean for supply chain design in the real world? My preliminary findings would

suggest that supply chain design can benefit from the type of collaboration that occurs organically in a virtual environment. Because there is a significant amount of transparency (i.e., creating content in the open where anyone can see it), collaboration occurs naturally. In a brick and mortar supply chain, product development, testing and prototyping can occur years before a consumer sees the product. Any mechanism that can shorten that time horizon will lead to better products that better meet the needs of their intended customers. Some firms have already shortened their supply chains by making use of Web 2.0 technology to allow people to design their own products...companies like Dell, Nike, and American Apparel. The compression in supply chain design is enhanced by the fact that consumers can see a three-dimensional version of the proposed product prior to its production.

The concept of vertical integration implies that a firm owns the factors of production from raw materials to product delivery. Second Life businesses tend to be vertically integrated. Does this mean that real life supply chains should be vertically integrated? It is uncertain at this juncture. However, the product development life cycle in Second Life is measured in days rather than months or years. Much can be learned from virtual entrepreneurs in this regard.

SUMMARY

This paper has presented preliminary findings of a study of virtual supply chains. I began with the question of whether supply chains exist in virtual worlds. While we found that they do, they are much shorter and compressed; supplier selection is simpler; and vertical integration is an acceptable part of owning a business in a virtual world.

REFERENCES

- [1] Hemp, P. (2006). "Avatar-Based Marketing." *Harvard Business Review* June 2006: 48-57.
- [2] Holger, W. (2007). "Second life, second identity?" *Journal of Targeting, Measurement and Analysis in Marketing* 15(3): 195.
- [3] Jonas, K. (2008). "Taking Care of Xerox Business-Virtually." *Research Technology Management* 51(1): 15.
- [4] Marcus, D. C. & B. Ray (2006). "Using Massively Multiplayer Online Role-Playing Games for Online Learning." *Distance Education* 27(2):187.
- [5] Paul, H. (2006). "Are you ready for E-tailing 2.0?" *Harvard Business Review* 84(10):28.
- [6] de Nood, D. & J. Attema (2006). "The Second Life of Virtual Reality." The Hague, EPN – Electronic Highway Platform.
- [7] Miles, M. B. & A. M. Huberman (1994). *Qualitative Analysis*, 2nd ed., Sage Publications, Thousand Oaks, CA.
- [8] Gibson, B. J., J. T. Mentzer, and R. L. Cook (2005). "Supply Chain Management: The Pursuit of A Consensus Definition." *Journal of Business Logistics* 26(2): 17-25.

**SUSTAINING SUPPLY CHAIN INTEGRATION: AN EXPLORATORY
INVESTIGATION**

Abirami Radhakrishnan**
507 MM Hall
Department of Information Science and Systems
Morgan State University
Baltimore, MD 21251
Phone: 443-885-1695
Email: aradhak@gmail.com

Sri V. Sridharan
119F Sistine Hall
Department of Management
Clemson University
Clemson, SC 29634
Phone: 864-656-2624
Email: suhas@clemson.edu

Douglas N. Hales
Department of Management
Ballentine Hall
University of Rhode Island
Kingston, RI 02881-0819
Phone: (401) 874-7882
E-mail: dhales@uri.edu

(All authors contributed equally for this paper)

**** Corresponding Author**

ABSTRACT

Supply Chain Integration is widely advocated as one of the important factors to attain superior supply chain performance. The evidence from studies examining the relationship between the level of integration and performance is mixed. While firms are able to achieve integration they find it is hard to sustain integration. In this paper we argue that to sustain integration, partner firms need to establish both integration capabilities and system wide performance evaluation capabilities. We draw on theories from a number of management disciplines to develop and propose a generalized conceptual framework for sustaining supply chain integration. We also develop and test our framework through a set of four research propositions. The results of our study show that supply chains that are highly integrated can sustain the integration by developing both integration and performance evaluation capabilities.

Key Words: Supply Chain Integration; Supply Chain Performance Evaluation; Sustainability, Theory Building

INTRODUCTION

It is generally accepted in the literature that successful manufacturing firms no longer compete head-to-head with other firms, but instead, they compete supply chain-to- supply chain. This means that firms rely on the complementation between their own internal competencies and those of their supply chain partners to achieve competitive advantage. To achieve complementary competencies among independently owned firms, a relatively new form of organizational governance, called Supply Chain Integration (SCI), is preferred over more traditional forms such as hierarchies or markets. The advantage of developing integrated supply chains is that it achieves the collaboration and communication capabilities of vertically integrated firms (i.e. hierarchies) while maintaining the flexibility and responsiveness capabilities of market-oriented governance structures. This allows firms with SCI capabilities to compete in new ways that help them, and their partner firms, achieve superior performance. Growing evidence suggests that the higher the degree of SCI capabilities achieved, the greater the potential benefits [1] [2]. Some studies show that SCI capabilities result in cost reductions, improvements in inventory, customer service, new product development, information and material flows, and financial performance for the focal firm (e.g. [1], [3]). However, these studies examined how SCI affects the performance of an individual focal firm rather than the performance of the supply chain as a whole.

Despite the aforementioned studies, which show a positive relationship between SCI and focal firm performance, others show that the positive relationship diminishes over time. Stated differently, firms that initially achieved SCI were unable to sustain it, and its benefits, in the long run (e.g., [4] [5]). The capabilities appear to have deteriorated over time - leading to a tremendous waste of financial and temporal resources and consequently, harming the ability of the supply chain to stay competitive. This finding is counter-intuitive in that supply chains

achieving superior performance through SCI capabilities should be able to maintain the capabilities over time.

The purpose of this study is to provide theoretical and empirical support for explaining how certain managerial interventions such as development of integration capability and performance evaluation capability can help supply chain partners sustain SCI in the long term and maintain superior supply chain performance.

LITERATURE REVIEW AND THEORY DEVELOPMENT

The literature suggests that superior business performance is achieved through competitive advantage by developing business capabilities that provide value-added activities to an end-user in unique ways or at lower prices than competitors [6]. Providing value-added activities in unique and low-cost ways require anticipating customer demand and responding to changes in the demand faster than competitors [7]. Independent firms can leverage each others' capabilities to achieve competitive advantage for the entire supply chain. The literature refers to the leverage of capabilities in a supply chain as an outcome of SCI.

To explain how companies create capabilities, the Resource-Based View (RBV) theory says that firms must invest resources in improving business processes [8]. These include, but are not limited to, demand management, new product development, customer relationship management, inventory management, evaluation capabilities.

1.0 Supply Chain Evaluation Capabilities (SCEC)

In addressing the importance of evaluation capabilities, Lambert and Pohlen [5], argue that they are essential to sustaining SCI capabilities. They stress the importance of developing supply chain evaluation capabilities (SCEC), which they refer to as supply chain performance measurement systems, for sustaining SCI. SCEC is generally described as an evaluation system designed to measure system-wide performance of all firms in a supply chain. It differs from performance metrics designed to measure specific functions or activities of individual firms [4] [5]. The need for SCEC is explained by Transaction Cost Economics theory (TCE). Two important key tenets of TCE [9] [10] are "Opportunism" and "Transaction risk". Opportunism arises when human actors in the exchange relationship are guided by considerations of self-interest with guile. In a supply chain context, there are several opportunities for "opportunistic" behavior where one player tries to get the maximum benefits at the expense of other players. A transaction's risk includes the risk that other parties in the transaction will shirk their agreed upon responsibilities. In order to curb opportunistic behavior and minimize transaction risks, SCEC can monitor each player in the supply chain.

In developing evaluation capabilities, academic literature emphasize the importance of functional-based measures at the firm level, such as operational measures (e.g., product quality, number of on-time deliveries, etc.) and financial measures (e.g., measures based on cost) (e.g. [11]). In deference to the functional-based measures, Bullinger et al. [12] suggest a greater emphasis should be placed on process-based metrics at the firm level. However, other studies stress the importance of process-based measures beyond the firm level, to include multiple firms in a supply chain (e.g. [4] [5]). They suggest that performance evaluation beyond the firm-level is important because improvement in individual firm performance (metrics) does not necessarily lead to improvement for the whole supply chain. For example, one firm may compete on flexibility – measuring success by how well it responds to changing customer requirements – while another firm may compete primarily on efficiency – measuring success by the cost of its operations. This difference in competitive priorities can lead to inter-firm conflict that harms performance. Addressing this conflict, the frameworks of Fisher [13] and Lee [14] classify supply chains based on the extent of demand and supply uncertainties faced – i.e. functional, innovative, agile, and risk hedging. Lambert and Pohlen [5] argue that “the goal should not be to identify specific metrics, but to provide a framework that allows management to develop the best performance measures for their situation”. They refer to this framework as a supply chain performance evaluation system, which we refer to as SCEC. They argue that failure to develop a SCEC system leads to misaligned performance measures that do not relate to the strategic non-financial performance and effectiveness of the system, ill-defined or inappropriate performance metrics used to guide decisions, and lack of accountability throughout the system.

Practitioners’ literature relies on the SCOR (supply chain operations reference) model for developing SCEC [15]. This model, developed by the Supply Chain Council, is the first cross-industry framework for evaluating supply chain performance and management. The SCOR model features four levels of supply-chain management:

- a) Level 1 provides a broad definition of the plan, source, make, deliver process types, and is the point at which a company establishes its supply chain competitive objectives.
- b) Level 2 defines 26 core process categories that are possible components of a supply chain system. A company can configure both its actual and ideal supply chain by selecting from these core processes.
- c) Level 3 provides a company with the information it needs to plan and set goals successfully for its supply-chain improvements through detailed process element information for each level 2 category. Planning elements include process element definitions, diagnostic metrics, benchmarks, best practices, and system software capabilities to enable best practices.
- d) Level 4 focuses on implementation, when companies put specific supply-chain improvements into play. Since changes at level 4 are unique to each company, the specific elements of the level are not defined within the industry-standard model.

(For the rest of the paper, please contact the corresponding author)

REFERENCES

- [1] Narasimhan, R. and Jayaram, J. (1998). "Causal linkages in supply chain management: an exploratory study of North American manufacturing firms". *Decision Sciences*, 29, 3, 579 – 599.
- [2] Vickery, S. K., Jayaram, J., Droge, C. and Calantone, R. (2003). "The effects of an integrative supply chain strategy on customer service and financial performance: an analysis of direct versus indirect relationships". *Journal of Operations Management*, 21, 5, 523-539.
- [3] Frohlich, M.T., Westbrook, Roy. (2001). "Arcs of integration: an international study of supply chain strategies". *Journal of Operations Management*, 19, 185–200.
- [4] Holmberg, S. (2000). "A systems perspective on supply chain measurements". *Internal Journal of Physical Distribution and Logistics Management*, 30, 10, 847-868.
- [5] Lambert, D. M. and Pohlen, T. L. (2001). "Supply chain metrics". *The International Journal of Logistics Management*, 12, 1, 1-19.
- [6] Porter, M.E. (1991). "Towards a dynamic theory of strategy". *Strategic Management Journal*, 12, Winter, 95-117.
- [7] Stalk, G., Evans, P. and Schulman, L.E. (1992). "Competing on capabilities: the new rules of corporate strategy". *Harvard Business Review*, 70, 2, 57-69.
- [8] Stank, T.P., Davis, B.R. and Fugate, B.S. (2005). "A strategic framework for supply chain oriented logistics". *Journal of Business Logistics*, 26, 2, 27-45.
- [9] Williamson, O. E. (1975). *Markets and Hierarchies: Analysis And Antitrust Implications*. New York: Free Press.
- [10] Williamson, O.E. (1985). *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. London: Free Press.
- [11] Beamon, B. M. (1998). "Supply chain design and analysis: models and methods". *International Journal of Production Economics*, 55, 3, 281-294.

- [12] Bullinger, H., Kuhner, M. and Van Hoof, A. (2002). "Analyzing supply chain performance using a balanced measurement method". *International Journal of Production Research*, 40, 15, 3533-3543.
- [13] Fisher, M. L. (1997). "What is the right supply chain for your product?" *Harvard Business Review*, 75, 2, 105-117.
- [14] Lee, H. L. (2002). "Aligning supply chain strategies with product uncertainties". *California Management Review*, 44, 3, 105-119.
- [15] Stewart, G. (1997). "Supply-Chain operations reference model (SCOR): the first cross industry framework for integrated supply-chain management". *Logistics Information System*, 10, 2, 62-67.

DIVERSITY IN INFORMATION EXCHANGE: AN EMPIRICAL STUDY OF MANUFACTURERS AND THEIR SUPPLY CHAINS

Tobin E. Porterfield, College of Business and Economics, Towson University,
Towson, MD 21252, 410-704-2586, tporterfield@towson.edu

Keywords: information exchange, information diversity, electronic commerce, supply chain

INTRODUCTION

The exchange of information between firms within supply chains is of great interest to both researchers and practitioners. Supply chain literature recognizes the value of exchanging information to improve supply chain performance. This phenomenon has been specifically addressed in efforts to dampen the bullwhip effect (Cachon and Fisher 2000; Lee et al. 1997; Machuca and Barajas 2004; Steckel et al. 2004). Complimenting this research, information technology (IT) researchers recognize that using IT to exchange interfirm information within the context of supply chains provides additional benefits to the participants (Bakos and Brynjolfsson 1993; Mukhopadhyay and Kekre 2002; Mukhopadhyay et al. 1995). The benefits to firms include decreased inventory investment (Mukhopadhyay et al. 1995), improved customer service (Allen et al. 1992), and reduced shipment errors (Srinivasan et al. 1994).

By observing information exchange occurring through an IT-enabled channel, unique measures of information exchange can be developed. This study uses specific IT-based measures to capture information exchange volumes and to identify the types of information are exchanged.

HYPOTHESES

As firms interact with their trading partners, they have the opportunity to minimize the cost of exchanging information by leveraging technology. In an EDI-enabled environment, firms may exchange large volumes of information with their trading partners at minimal cost. Once the initial cost of formatting the information and establishing the communication link is made, the incremental cost of each additional document is minimal. Although the upfront cost of creating an EDI relationship has been noted as a deterrent to EDI implementation (Crum et al. 1998; Iacovou et al. 1995), these become sunk costs once the firm implements the technology. Once the electronic channel is established, firms can lower costs by increasing the volume of information exchanged electronically. Thus,

H1: Information exchange volume is positively associated with firm performance.

Information diversity is a measure of the unique types of information exchanged by a firm. A base level of information exchange must occur in order to do business in the supply chain. At a minimum, the buyer must identify what they want to purchase and when they require delivery. The supplier then confirms the pricing and availability back to the buyer. When the transfer of the physical product is complete, an invoice and payment must be exchanged. But beyond basic transactions, firms can create unique competitive advantages by exchanging additional information (Dyer and Singh 1998). Thus,

H2: Information exchange diversity is positively associated with firm performance.

METHODOLOGY AND DATA

Data Collection

Data for the measurement of B2B information exchange is gathered from an electronically mediated industrial exchange network. This proprietary data has been made available by one of the world's largest providers of B2B integration services. Additional data on firm characteristics and performance are gathered from Standard and Poor's Compustat database.

TABLE 1: MEASURES AND DATA

Variable	Definition
Dependent Variable	
<i>INVENTORY_TURNOVER_i</i>	The ratio of a firm's cost of goods sold to the firm's inventory value. Average for two years.
Independent Variables	
<i>INFO_VOLUME_i</i>	The average number of information exchanges with trading partners through an electronic intermediary. The average is based on quarterly observations averaged across two years of data.
<i>INFO_DIVERSITY_{it}</i>	The average number of unique information types exchanged through an electronic intermediary. The average is based on quarterly observations averaged across two years of data.
Control Variables	
<i>FIRM_SIZE_i</i>	Firm average quarterly sales. The average is calculated across two years of data.

Sample

Firms were identified for inclusion in the study based on their level of participation in the host EDI exchange network. A manufacturing focus was adopted to minimize the variance created by including multiple echelons of the supply chain. The resulting twenty-three firms are included in the sample.

RESULTS AND DISCUSSION

Full results from the OLS regression are provided in Table 2.

TABLE 2: COEFFICIENT ESTIMATES FROM OLS REGRESSION

INVENTORY_TURNOVER	Coef. (Std Err)	P> t	sig
<i>Explanatory Variables</i>			
(log)INFO_VOLUME	0.8024 (0.1862)	0.000	**
(log)INFO_DIVERSITY	-1.2356 (0.4715)	0.017	**
constant	-1.868 (2.0696)	0.378	ns
<i>Control Variables</i>			
(log)FIRM_SIZE	-0.3023 (0.2172)	0.378	ns

**<.01 *<.05 +<.1 significance level

The positive result for Hypothesis 1 supports the results of prior literature on the relationship between information volume and firm performance. Although not surprising, it is useful to address the issue using unique quantitative measures from an electronically-mediated exchange network.

The surprising result for Hypothesis 2 requires further analysis. The negative and statistically significant result is important. For firms in this sample, exchanging more types of information with trading partners has a detrimental effect on inventory turnover. Since this sample focuses on manufactures, this result may indicate that increased information diversity in manufacturing supply chains allows trading partners to act opportunistically. This situation may be affected by the level of asset specificity of products transacted by manufacturers. An alternative explanation of this surprising result has been identified in the literature. It has been suggested that as firms take advantage of the technology which allows them to exchange more types of business

documents, there is no guarantee that the firms are able to integrate that information into their internal or external processes (Clarke 1992; Massetti and Zmud 1996). If the data is exchanged but is not actually used in the decision making process then it adds no value. This situation then may be unique to the manufacturing sector. Other researchers using a simulation methodology have noted that information overload can create detrimental effects (Steckel et al. 2004). This suggests that the additional information types exchanged are not only failing to contribute to performance but may actually be detrimental to performance by causing noise in the channel.

REFERENCES

- Bakos, Yannis, and Erik Brynjolfsson. 1993. "From vendors to partners: Information technologies and incomplete contracts in buyer seller relationships." *Journal of Organizational Computing* 3:301-329.
- Cachon, Gerard, and Marshall Fisher. 2000. "Supply chain inventory management and the value of shared information." *Management Science* 46:1032-1048.
- Cachon, Gerard P., and Martin A. Lariviere. 2001. "Contracting to assure supply: how to share demand forecasts in a supply chain." *Management Science* 47:629-646.
- Clarke, R. 1992. "A contingency model of EDI's impact on industry sectors." *Journal of Strategic Information Systems* 1:143-151.
- Crum, Michael R., Deborah A. Johnson, and Benjamin J. Allen. 1998. "A longitudinal assessment of EDI use in the U.S. motor carrier industry." *Transportation Journal* 38:15.
- Dyer, Jeffery H, and Harbir Singh. 1998. "The relational view: Cooperative strategy and sources of interorganizational competitive advantage." *Academy of Management Review* 23:660-679.
- Iacovou, Charalambos L., Izak Benbasat, and Albert S. Dexter. 1995. "Electronic data interchange and small organizations: Adoption and impact of technology." *MIS Quarterly* 19:465-485.
- Lee, Hau L., V. Padmanabhan, and Seungjin Whang. 1997. "Information distortion in a supply chain: The bullwhip effect." *Management Science* 43:546-558.
- Machuca, Jose A.D., and Rafael P. Barajas. 2004. "The impact of electronic data interchange on reducing bullwhip effect and supply chain inventory cost." *Transportation Research Part E* 40:209-228.
- Massetti, Brenda, and Robert W. Zmud. 1996. "Measuring the extent of EDI usage in complex organizations: Strategies and illustrative examples." *MIS Quarterly* 20:331-345.
- Mukhopadhyay, Tridas, and Sunder Kekre. 2002. "Strategic and operational benefits of electronic integration in B2B procurement processes." *Management Science* 48:1301-1313.
- Mukhopadhyay, Tridas, Sunder Kekre, and Suresh Kalathur. 1995. "Business value of information technology: A study of electronic data interchange." *MIS Quarterly* 19:137-156.

- Srinivasan, Kannan, Sunder Kekre, and Tridas Mukhopadhyay. 1994. "Impact of electronic data interchange technology on JIT shipments." *Management Science* 40:1291-1304.
- Steckel, Joel H., Sunil Gupta, and Anirvan Banerji. 2004. "Supply chain decision making: Will shorter cycle times and shared point-of-sale information necessarily help?" *Management Science* 50:458-464.

A PRODUCT RECOVERY AND LOGISTICS NETWORK DESIGN MODEL

Nasr-Eddine Dahel, Fisher Graduate School of International Business, Monterey Institute of International Studies, 460 Pierce Street, Monterey, CA 93940, 831-647-4602, edahel@miis.edu

ABSTRACT

For a product recovery operation to function effectively the existing logistics network of facilities must be redesigned in the form of a closed-loop system to handle the arising forward flows of goods to the customer and return flows of cores from end users. At the same time the organization must plan the quantities to collect, disassemble, store and remanufacture in such a way to meet demand at the minimum cost. This paper presents a multi-period cost minimization mixed integer programming model that simultaneously solves for the location of the remanufacturing/distribution and disassembly facilities, the transshipment, production, stocking, and disposition of the optimal quantities of remanufactured products and cores.

Keywords: Closed-Loop Supply Chains, Product Recovery, Reverse Logistics, Network Design

INTRODUCTION

Remanufacturing is gaining justifiable popularity among society, government, and industry worldwide. Until recently much of the remanufacturing was driven by cost cutting considerations and limited to low-volume high-value items. However many companies, attempting to combine good business sense with environmental sustainability, are now increasingly remanufacturing high-volume low-value items such as single-use cameras, mobile phones, ink-jet printers, and cartridges (Guide et al 2003). For instance, Eastman Kodak Company reuses on average 76% of the weight of a disposed camera in the production of a new one (Savaskan et al 2004).

Remanufactured products are generally upgraded to the quality standards of new products so that they can be sold as new products. The production and distribution systems which combine product recovery and remanufacturing are referred to as closed-loop supply chains. 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. The intercession role that closed-loop networks play between these two heterogeneous markets gives rise to the issue of coordination between supply and demand in a recovery operation. 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 (Fleischmann et al 2001).

Another major characteristic of 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. Moreover, even if technically feasible, a recovery operation may not be economically attractive. Since total recovery costs (collection, disassembly, processing, and transportation) depend to a large extent on the structure of the logistics network (the relative location and size of disassembly centers to plants, collection points and disposal sites, and on the relative location and size of plants to markets, then optimal design of the closed-loop network becomes critical to the economic viability of the recovery operation.

For a recovery operation to function effectively, the issues of: (1) mismatches between supply and demand, (2) quality uncertainty, and (3) network structure need to be taken into account when formulating closed-loop logistics models.

Remanufacturing can be carried out by a local manufacturer or an original equipment manufacturer (Pranab and Harry 2001). In this paper we consider 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 the most economical way.

The objective of the paper is to formulate a multi-period cost-minimization integrated network design and recovery planning model that provides unambiguous answers to such questions as: Which plants and which disassembly centers should be open and operated during the planning horizon? Which plants should service which market's product demand, and in what quantity in every period? How many units of used products are to be collected from each customer zone and shipped to each disassembly center in every period? How many reusable units should each disassembly center ship to each plant in every time period? How many reusable units should be held in inventory at each disassembly center in every period? In which disposal site, and in what quantity per time period should non-reusable units be disposed of?

MODEL DEVELOPMENT

The proposed integrated network design and recovery planning model follows the closed loop network structure shown in Fig. 1. The supply chain includes two types of flows: forward and reverse flows. Forward flows represent transportation of remanufactured products from plants to retailers. The reverse flows represent: (1) transportation of used products from retailers to

disassembly centers, (2) transportation of used non-reusable units from disassembly centers to disposal sites, and (3) transportation of reusable units from disassembly centers to plants.

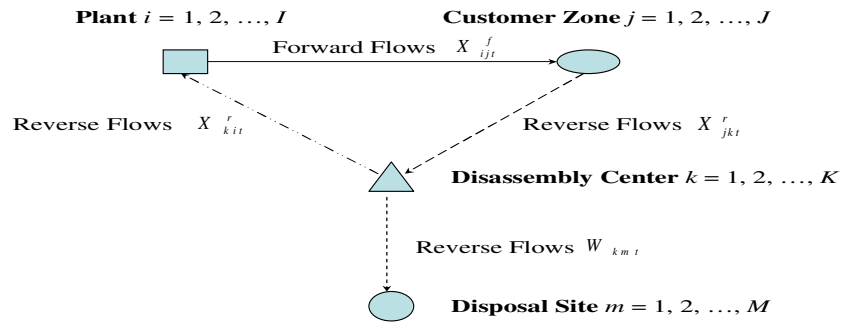


Fig. 1 Closed-Loop Network

Assumptions

In order to facilitate the formulation of the model, the following assumptions are postulated. The supply chain facilities (plants, customer zones, disassembly centers, and disposal sites), already exist. Product demand and return at customer zones are known. Customer zones are responsible for collecting used products from consumers. Plant production capacities, product collection capacities, and disassembly center capacities are known. There is a given recovery ratio of used products into reusable units. Inventory of reusable units is held at the disassembly centers. 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 zones or demand points, $j \in J$;
- M = Set of disposal sites, $m \in M$;
- T = Set of time periods, $t \in T$;

Supply/Demand Data

- D_{jt} = Product demand at customer zone j during period t ;
- S_i = Plant i production capacity per period;
- S_j = Customer zone j used product collection capacity per period;
- S_k = Center k disassembly capacity per period;
- θ_k = Center k inventory storage capacity per period;
- A_k = Center k minimum output shipment to plants per period;
- d_{jt} = Product return forecast at customer zone j during period t ;
- λ = Product recovery ratio during disassembly;

Cost Data

F_i = Fixed cost of opening and operating plant i ;

F_k = Fixed cost of opening and operating disassembly center k ;

C_{ki} = Per unit remanufacturing cost at plant i using a reusable unit sourced from center k .
This cost includes production cost at plant i , transportation cost from k to i , and cost of material from center k ;

C_{jk} = Per unit disassembly cost at center k of a used product collected in customer zone j .
This cost includes collection cost at zone j , transportation cost from j to k , and disassembly cost at k ;

C_{km} = Per unit disposal cost at site m of a non-reusable unit processed at center k . This cost includes disposal cost at site m and transportation cost from k to m ;

H_k = Per unit per period inventory holding cost of a reusable unit in inventory at center k ;

P_j = Unit penalty cost for not collecting a used product returned in customer zone j ;

G_j = Unit penalty cost of not serving demand of customer zone j . Observe that G_j could be quantified by taking the relative importance of the different customer zones j 's into account; alternatively it could be related to the cost of meeting demand of zone j by resorting to external suppliers.

T_{ij} = Unit transportation cost of a product from plant i to customer zone j ;

Variables

X_{ijt}^f = Forward flow: units shipped from plant i to customer zone j in period t ;

X_{jkt}^r = Reverse flow: units of used product shipped from customer zone j and to center k in period t ;

X_{kit}^r = Reverse flow: reusable units shipped from center k to plant i in period t . Observe that since no inventory of remanufactured units is held at the plants, this quantity also reflects the number of units produced in plant i out of cores sourced from center k in period t ;

W_{kmt} = Non-reusable units shipped from disassembly center k to site m for disposal in period t ;

B_{jt} = Units of unsatisfied demand from customer zone j in period t ;

I_{kt} = Disassembled units held in inventory at center k at the end of period t ;

U_{jt} = Uncollected units of returns from customer zone j at the end of period t ;

$Z_i = \begin{cases} 1, & \text{if product is produced in plant } i; \\ 0, & \text{otherwise.} \end{cases}$

$Y_k = \begin{cases} 1, & \text{if used product is disassembled in center } k; \\ 0, & \text{otherwise.} \end{cases}$

Model Formulation

$$\begin{aligned} \text{Min } & \sum_{i \in I} F_i Z_i + \sum_{k \in K} F_k Y_k + \sum_{t \in T} [\sum_{j \in J} (P_j U_{jt} + G_j B_{jt}) + \sum_{k \in K} H_k I_{kt} + \\ & \sum_{i \in I} \sum_{k \in K} C_{ki} X_{kit}^r + \sum_{j \in J} \sum_{k \in K} C_{jk} X_{jkt}^r + \sum_{k \in K} \sum_{m \in M} C_{km} W_{kmt} + \sum_{i \in I} \sum_{j \in J} T_{ij} X_{ijt}^f] \end{aligned} \quad (1)$$

Subject to:

$$\sum_{j \in J} X_{ijt}^f \leq S_i Z_i, \quad i \in I, t \in T; \quad (2)$$

$$\sum_{i \in I} X_{ijt}^f + B_{jt} = D_{jt}, \quad j \in J, t \in T; \quad (3)$$

$$\sum_{k \in K} X_{kit}^r = \sum_{j \in J} X_{ijt}^f, \quad i \in I, t \in T; \quad (4)$$

$$I_{k,t-1} + \lambda \sum_{j \in J} X_{jkt}^r = \sum_{i \in I} X_{kit}^r + I_{kt}, \quad k \in K, t \in T; \quad (5)$$

$$I_{kt} \leq \theta_k Y_k, \quad k \in K, t \in T; \quad (6)$$

$$\sum_{k \in K} X_{jkt}^r + U_{jt} = d_{jt}, \quad j \in J, t \in T; \quad (7)$$

$$\sum_{k \in K} X_{jkt}^r \leq S_j, \quad j \in J, t \in T; \quad (8)$$

$$\sum_{j \in J} X_{jkt}^r \leq S_k Y_k, \quad k \in K, t \in T; \quad (9)$$

$$\sum_{i \in I} X_{kit}^r \geq A_k Y_k, \quad k \in K, t \in T; \quad (10)$$

$$\sum_{m \in M} W_{kmt} = (1 - \lambda) \sum_{j \in J} X_{jkt}^r, \quad k \in K, t \in T; \quad (11)$$

$$Y_k, Z_i = \{0, 1\}, \quad i \in I, k \in K; \quad (12)$$

$$X_{ijt}^f, X_{jkt}^r, X_{kit}^r, W_{kmt}, I_{kt}, B_{jt}, U_{jt} \geq 0, \quad i \in I, j \in J, k \in K, m \in M, t \in T. \quad (13)$$

Constraints (2) state 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 customer zone j , and demand requirement for that zone at time period t , and account for the possibility of unsatisfied demand at that zone. Unsatisfied demand occurs when not enough used units are collected or when product demand is greater than production and/or disassembly capacities. Equation (4) is a material balance constraint ensuring that the total number of reusable units going into a plant i (or reverse flow) equals the total number of remanufactured units coming out of that plant (or forward flow) in every time period. Input into plant i can be sourced from any open disassembly center k and the output of such a plant can be shipped to any customer zone j . Constraints (5) ensure product flow balance between inventory of reusable units, processing of used units, and shipment of reusable units at disassembly center k in time period t . Used units are assumed to yield λ reusable units. Inventory at center k 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 reusable 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 used units, forecasted return of used units and uncollected units of used product at customer zone j in time period t . Observe that the number of used units collected at zone j determines the total reverse flow from zone j to all open disassembly centers. Constraints (8) require that the total number of used units collected at customer zone j in time period t to be less than the collection capacity of that zone. 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. Constraints (10) require that the total flow out of any open center k in period t meets the minimum output requirement A_k for that center. Management may specify A_k in such a way to ensure a minimum outflow activity for center k thereby preventing such a center from becoming just an inventory storage location of reusable units. Constraints (11) specify the number of non-reusable units transported from disassembly center k to site m for disposal. Observe that a proportion $(1-\lambda)$ of the used units collected from all customer zones and processed in center k are non-reusable and must therefore be disposed of.

The objective function (1) minimizes the total multi-period cost of remanufacturing, collection, disassembly, disposal, inventory, and transportation of the recovery operation. The components of the objective function may be described as follows: Fixed cost at plants and disassembly centers over the entire planning horizon $= \sum_{i \in I} F_i Z_i + \sum_{k \in K} F_k Y_k$. Multi-period variable production cost at the plants $= \sum_{t \in T} \sum_{i \in I} \sum_{k \in K} C_{ki} X_{kit}^r$. Inventory costs at processing centers $= \sum_{t \in T} \sum_{k \in K} H_k I_{kt}$. Penalty cost of unsatisfied demand at customer zones $= \sum_{t \in T} \sum_{j \in J} G_j B_{jt}$. Penalty cost of uncollected returns at customer zones $= \sum_{t \in T} \sum_{j \in J} P_j U_{jt}$. Collection, transportation, and processing costs of used units $= \sum_{t \in T} \sum_{j \in J} \sum_{k \in K} C_{jk} X_{jkt}^r$. Disposal and transportation costs of non-reusable units $= \sum_{t \in T} \sum_{k \in K} \sum_{m \in M} C_{km} W_{kmt}$. Transportation cost of new units from plants to retailers $= \sum_{t \in T} \sum_{i \in I} \sum_{j \in J} T_{ij} X_{ijt}^f$.

CONCLUSION

Collection, recovery, remanufacturing, transportation and distribution are complicated decisions subject to structural constraints of the logistics network as well as environmental factors concerning supply and quality uncertainty of the used products. The proposed model is able to effectively capture these constraints and determine the optimal design of the recovery network, and specify the optimal flows, along with the optimal levels of unsatisfied demand and collection of used units over a multi-period planning horizon in such a way to minimize the total cost of the recovery operation. The proposed mixed-integer programming model is computationally efficient – no more than 1.5 CPU seconds were required to solve a problem with 10 potential plants, 10 customer zones, 5 potential disassembly, and 3 disposal sites – and hence capable of providing optimal solutions to realistic size recovery operations.

Full paper and references available upon request.

AGENCY THEORY AND THE SUBOPTIMAL ALLOCATION OF SCARCE DETAIL PARTS WITHIN A MANUFACTURING FIRM

Thomas A. Schuessler
10705 Winding River Drive, Bellevue, NE 68123
Schuta53@gmail.com
(402) 238-7295

Robert A. Marose: St. Joseph's College: Department of Business
32 Audubon Avenue, Patchogue, NY 11772
Ram3840@optonline.net
(516) 799-5009

Linda E. Castro: Automatic Data Processing
Walt Whitman Road, Melville, NY 11111
linda_e_castro@adp.com
(631) 753-8671

Abstract

The general assumption is that managers are supposed to act in ways that increase shareholder wealth. However, Agency Theory states that "... when the interests of managers diverge from those of the owners, then the managers' decisions are more likely to reflect the managers' preferences than the owners' preferences" (Pearce 43). In contrast to the maximization of shareholder wealth managers often engage in activities that feather their own nests (Wheelen 248). This paper demonstrates that managers in small firms will often engage in actions that benefits their department at the expense of the firm as a whole (Wheelen 247)

Literature Review

As previously stated, Agency Theory contends that due to the separation of ownership and management there can be times when the goals and desires of the managers differ from the goals and desires of the owners. Due to this divergence of goals owners incur costs associated with having hired managers. "These costs may occur in explicit ways, such as the excessive use of perquisites or, in implicit ways such as sub-optimal decision making" (Oswald 1991). A disproportionate percentage of the literature concerning Agency Theory and companies is related to the conflicts between owners and managers in large publicly traded firms (Conlon & Parks 1988; Eisenhardt 1985; Amihud and Lev 1981; Fama and Jensen 1983). Much of this literature focuses on attacking the problem by increasing the ownership stake of the senior managers (Oswald 1991; Argawal & Mandelker 1987).

Information Asymmetry is described as one of the key factors in Agency Theory. Since managers usually have more information about the company and its activities, it becomes difficult for the owners to verify whether the actions of the managers are designed to maximize the owners' utility or the managers' utility (Husted 2007). "When the principal has information to verify agent behavior, the agent is more likely to behave in the interests of the principal" (Eisenhardt 1989).

The amount of literature on Agency Theory and private firms is disproportionately small. "Private firms have received relatively little theoretical and empirical attention despite their importance in world-wide wealth creation and business formation" (Durand & Vargas 2003). Difficulty in obtaining information on privately owned firms makes it more difficult to study them in comparison to publicly traded firms.

The basic solution to the agency problem is a combination of monitoring and incentives to ensure that the managers act in ways that increase shareholder wealth. The incentives should align managers' utility with owners' utility. Results show that less conscientious individuals increase effort through incentive alignment and monitoring. However, less conscientious individuals increase task performance through incentive alignment, but not through monitoring (Fong & Tosi 2007).

In addition, teams of executives from across divisions or units can be formed to focus performance on organizational goals rather than personal goals (Pearce 45).

One criticism of Agency Theory is that it is "hardly subject to empirical test since it rarely tries to explain actual events" (Perrow 224). We, however, will be explaining actual events and providing an example to illustrate the problem.

The Case

Overview of the Company: The firm in this case is a manufacturing firm located on Long Island, New York, which employs about three hundred people. Its products are assembled from a combination of purchased detail parts and in-house-made detail parts. The company suffers from frequent shortages of both purchased and in-house-made parts. The materials group is responsible for the purchase of the parts that are obtained from vendors as well as the purchase of the raw materials that are machined by the company. This group is also responsible for the flow of in-house machining and the timely arrival of parts at the assembly department.

There are three principal final uses for the company's detail parts. The first is to be used as components for the products that are assembled and sold to the firm's direct customers. The second is to be sold to customers as spare parts. This allows the customers and independent repair stations to use these detail parts to repair the units at their facilities. The final use is for repair of customer owned units by the manufacturer. This is always the most profitable use for this firm.

The Agency Problem: The materials group is judged on two out of the three possibilities above. It has revenue targets for both production sales and sales of spare detail parts to the customer. The materials group manager’s performance reviews and any bonuses are based on meeting these targets. They have, therefore, a strong temptation to allocate scarce detail parts to those uses even when it is more profitable to use them in repairs or overhauls, which are significantly more profitable.

A specific example occurred when housings for a particular type of unit were very delinquent from the machine shop. The repair arm of the company had orders due in the month that the machining of twenty of these housings was completed. There were six customer-owned units awaiting repair that required housing replacement. There were also orders for these housings as spare parts. The spares order, however, was not due until the following month. Figure 1 shows these orders.

Figure 1: Open orders for the Housings in Question.

Open Orders	Month 1 (Housings Available in this month.)	Month 2	Month 3
Repair Division	20	0	20
Spares Orders	0	20	0

As Figure 1 indicates the repair division had an open order for 20 of these housings. However, the materials group allocated them to cover the spares orders that were due in the following month, by shipping them in month 1 to cover a shortfall in their spare parts revenue.

This hurt the company in two ways. The housings were sold as spares, which yielded significantly less profit in that month. More importantly, this action makes it difficult to increase repair and overhaul revenues. Customers will repair the units themselves or send them to independent shops for repair, rather than wait for an extended period. This will ensure that future spares orders will be received as these independent shops replenish their inventory. This is a very suboptimal result for the company since the rich profits from the repairs will be enjoyed by these independent shops at the expense of the manufacturing firm as a whole. The following equations illustrate this:

Equation 1 Profit from spares

$$S - C = Sp$$

Where: S = Sales price per unit
 C = Accumulated cost per unit
 Sp = Unit profit from spares sales

$$\$4000 - \$2000 = \$2000$$

Equation 2 Profit from repairs

$$R - C - L = R_p$$

Where: R = Unit price of repair
 C = Accumulated Costs
 L = Labor and overhead associated with repair
 R_p = Profit from repairs

$$\$8500 - \$2000 - \$2000 = \$4500$$

These two equations illustrate just one example of the suboptimization that occurs on a regular basis at the firm.

Addressing the Problem: The problem in many firms is a misalignment of incentives for managers. In this case the managers in the materials group have their performance judged on meeting revenue targets for sales of new units and sales of spare parts. They have no incentive to determine whether allocating the parts to the repair arm of the firm will yield higher profits for the company as a whole. Their performance reviews, salary increases, and bonuses should be based on the allocation of scarce detail parts in a manner that will most increase shareholder wealth most.

Since the goal is to produce the most revenues and profits in a given period, the managers' actions should not result in an all-or-nothing strategy for allocating scarce parts. In Figure 1 the repair arm of the company was due twenty housings in the first month. At the time it had only six units awaiting the housings for repair. The remaining housings could have been allocated to spares sales. The optimal distribution is demonstrated in Equation 3.

Equation 3
$$(Q_r * P_r) + ((Q - Q_r) * P_s)$$

Where: Q = Quantity of housings available
 Q_r = Quantity required for repairs
 P_r = Profit on repairs
 P_s = Profit on spares sales

$$(6 * 4500) + ((20 - 6) * 2000) = \$55000$$

This formula yields a profit for the company in the amount of \$55,000. When the twenty housings were sold for spares at \$2000 each, they only produced a profit of \$40,000.

However, in the given example, the next production lot of housings was not expected to be completed until Month three in the example. This meant that if any in-house repairs of these units were expected in Month 2 there would again be units that could not be repaired in a timely manner. Therefore, a forecast was made to determine the likelihood that housings would be required by the repair division in Month 2. The estimate was that two housings would be needed. Equation 4 demonstrates the optimal distribution of parts with an estimate of future requirements.

Equation 4

$$(Q_r * P_r) + (F_r * P_r) + ((Q - (Q_r + F_r)) * P_s)$$

Where F_r = Forecast need until replenishment

$$(6 * 4500) + (2 * 4500) + ((20 - (6 + 2)) * 2000) = 60,000$$

This is the optimal distribution for this example since it allocates the housings to the most profitable use while still allowing the materials group to ship twelve of the housings a month early to help alleviate the expected revenue shortfall.

Summary and Comments

This case clearly demonstrates that Agency Theory can be applied to explain actual events despite criticism to the contrary. It also demonstrates that agency problems can occur in middle management, whereas the majority of literature has focused on senior management.

The application of Agency Theory to a specific problem involving middle management requires special attention to the alignment of owners' utility and managers' utility. In the case of senior management in large publicly traded companies, the usual solution to the misalignment of owner and management utility is to require that the senior managers own a significant number of shares in the firm. Since this makes them owners, they are more likely to engage in strategies and activities that will increase shareholder wealth. Since the company is privately-owned another solution must be found.

The critical factors in the case are information asymmetry and misalignment of incentives. The materials group is required to report its revenues for sales of production units and sales of spare parts to ascertain that it has met or exceeded its goals for the period. The repair arm of the firm reports its revenues separately to determine if they have met or exceeded their revenue goals. The information is asymmetrical in that senior management and the owner never see how the allocation of scarce detail parts affect total revenues and profits for the company as a whole. Senior management is usually satisfied if all revenue and profit targets are met. There is no consideration of whether higher revenues and profits could have been attained if parts had been allocated differently. Additionally, the materials managers have no incentive to allocate scarce parts more profitably because their performance is judged to be acceptable since they have met their targets.

To solve this problem senior management must align the middle managers' incentives with the desired actions, which are those that maximize shareholder wealth. Specifically, their performance reviews, salary increases, and bonuses should be based on the optimal allocation of detail parts in addition to other important criteria.

To ensure that this occurs, a committee of senior managers should be formed to design a new reporting system that will not only measure performance of each unit of the firm, but

will also assure that the actions of the middle managers have maximized shareholder wealth.

Small privately-owned firms are not exempt from agency problems, nor are agency problems confined to senior management. The owners of such firms should actively review the actions of the managers to ensure that the managers' actions are producing the maximum value for the owners. This review should cover every phase of the company's operation and every level of management in the company.

Works Cited

- [1] Amihud, Y and Lev, B. 1981 Risk Reduction as A Managerial Motive For Conglomerate Mergers Bell Journal of Economics, 12: 605-617
- [2] Argawal, A. and Mandelker, G. 1987 Managerial Incentives and Corporate and corporate Investment and Financing Decisions. Journal of Finance, 42, 823-837
- [3] Durand, R. and Vargas, V. 2003 Ownership, Organization, and Private Firms' Efficient Use of Resources. Strategic Management Journal Vol. 24, Issue 7, p. 667
- [4] Eisenhardt, K.M., 1985 Control, Organization, and Economic Approaches Management Science, 31 134-149
- [5] Fama, E.F. And Jensen, M.C. 1983 Separation of Ownership and Control Journal of Law and Economics 26, 301-325
- [6] Fong, E.A. and Tosi, H.L. Jr., 2007 Effort, Performance, and Conscientiousness: An Agency Theory Perspective Journal of Management 33, no. 2, 161-179
- [7] Conlon, E. and Parks, J. 1988 The Effects of Monitoring and Tradition On Compensation Arrangements: An Experiment On Principal/Agent Dyads Best Papers Proceedings pp. 191-195 Anaheim, CA: Academy of Management
- [8] Husted, B.W. 2007 Agency Information And The Structure Of Moral Problems In Business Organization Studies Vol. 28, Iss. 2, p. 177
- [9] Oswald, S. and Jahera, J., Jr., 1991 Research Notes And Communications: The Influence Of Ownership On Performance: An Empirical Study Strategic Management Journal, 12 321-326
- [10] Pearce, J.A. II., and Robinson, R.B. Jr., 2005 Strategic Management: Formulation, Implementation, and Control New York: McGraw-Hill
- [11] Perrow, C. 1986 Complex Organizations: A Critical Essay New York: McGraw-Hill
- [12] Wheelen, T.L. and Hunger, J.D. 2000 Strategic Management & Business Policy Upper Saddle River: Prentice Hall