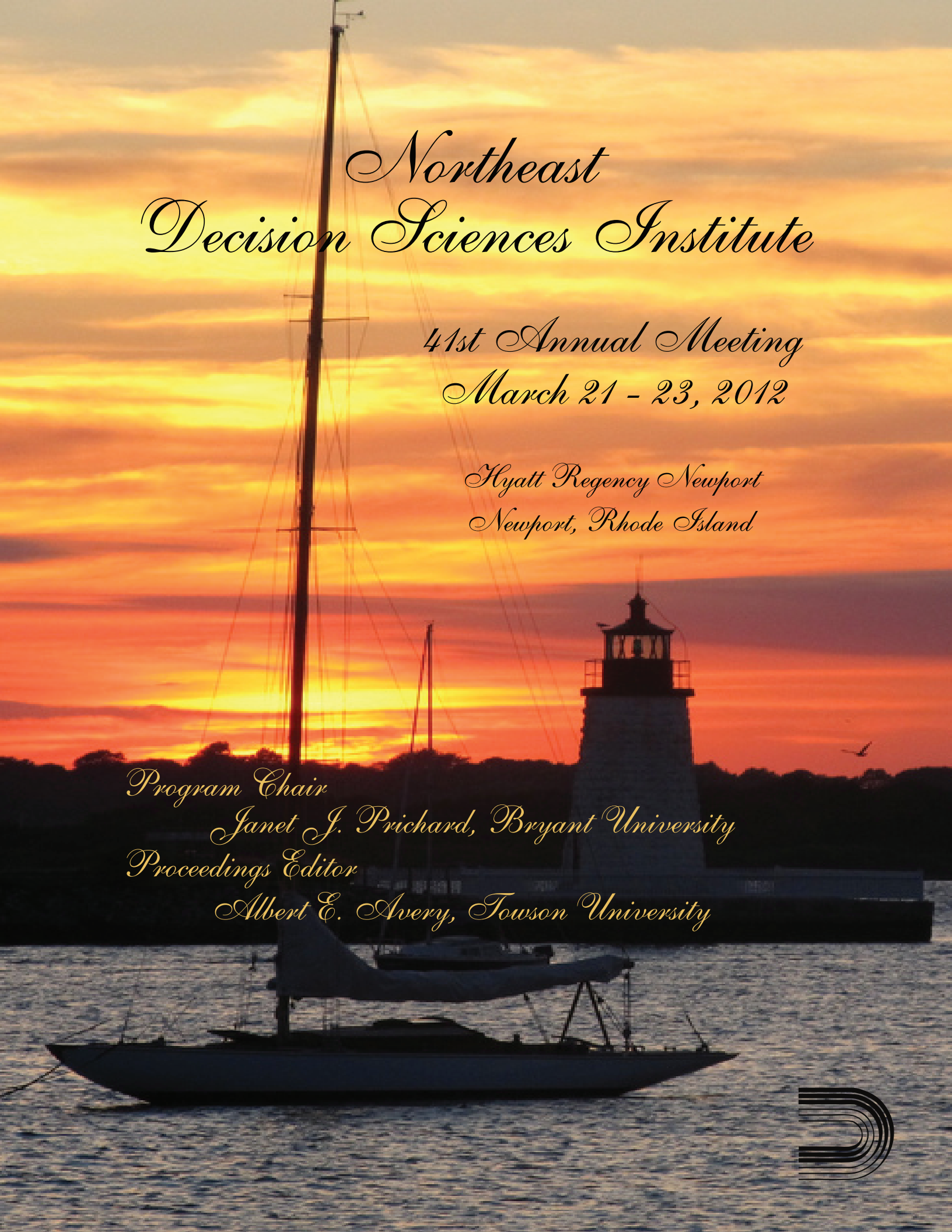


Northeast Decision Sciences Institute

41st Annual Meeting
March 21 - 23, 2012

Hyatt Regency Newport
Newport, Rhode Island

Program Chair
Janet J. Prichard, Bryant University
Proceedings Editor
Albert E. Avery, Towson University

INTER-METHOD RELIABILITY OF CRITICAL THINKING CONSTRUCTS IN THE INTRODUCTORY ACCOUNTING COURSES

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

ABSTRACT

The objective of this research-in-progress is to examine the inter-method reliability of multiple critical thinking measurements used in an introductory accounting course. Students in an introductory course in accounting will be given two separate measurements on critical thinking skills. One method of measurement will consist of a battery of objective questions using a multiple choice format. A second method of measurement will use open-ended questions where the students responses will be evaluated using a rubric to assess their critical thinking skills.

Keywords: accounting education, critical thinking, curricula

LITERATURE REVIEW

Competency in critical thinking skills is one of the core requirements of the accounting profession as espoused by the American Institute of Certified Public Accountants (AICPA). Indeed, the AICPA Core Competency Framework for Entry into the Accounting Profession (EAP) [2] promotes a list of core competencies required for students to successfully enter the accounting profession. Critical thinking is one of the broad business perspective category items.

Conveniently, various introductory accounting textbooks require students to demonstrate competence in applying "critical thinking skills." These textbooks have coded, or categorized, various questions, exercises, and problems, including both homework assignments and instructors text banks, as having a critical thinking component. Our research sets out to determine the inter-method reliability of some of these measures and whether they can be used interchangeably to measure critical thinking skills in solving accounting problems.

As indicated by Barel et.al. [3], there is no universal definition of what constitutes critical thinking skills. Perhaps the core of the definition is the ability to solve unstructured problems. Arum and Roksa [1] have reviewed several authoritative sources on the importance of acquiring critical thinking skills during the process of completing an under graduate education. In their review, they cite Derek Bok, the former president of Harvard University on the unanimous consensus of faculty members agreeing that critical thinking skills should be the principal objective of a college education. They also cite the American Association of University professors (AAUP) position that "critical thinking is the hallmark of American education-an education designed to create thinking citizens for a free society."
The following description of critical thinking is provided by the AICPA EAP [2] framework database:

> Critical thinking encompasses the ability to link data, knowledge, and insight together from various disciplines to provide information for decision making. Being in tune with the "big picture" perspective is a necessary component for success. Individuals entering the accounting profession should be able to communicate to others the vision, strategy, goal, and culture of organizations.

In addition, universities and colleges that are accredited by AACSB are required to complete schedule assessments of the goals and objectives that are delineated within their curriculum. Critical thinking skills are usually a major area that is assessed across various courses within the curriculum. The AACSB provides seminars which inform business schools of how to organizationally implement a critical thinking approach in a curriculum.

Various introductory accounting textbooks provide problems that require students to demonstrate competence in employing "critical thinking" skills. These textbooks have included supplementary instructor resources with various assessment instruments coded as critical thinking skills competencies. These instruments include questions, exercises, and problems, both homework assignments as well as in the instructor's test banks. In adopting an introductory accounting textbook, academics may be selecting assignments and exam questions for their students that have been coded as measuring critical thinking skills. Some business schools may want to use these measurement instruments to assess a critical thinking competency as part of their school's AACSB Assurance of Learning (AoL) program objectives.

The purpose of this research study is to determine the inter-method reliability of the measurements classified as measuring critical thinking skills in an introductory accounting textbook. Stated simply, are these instruments that have been coded as measurements of critical thinking skills reliably measuring this competency?

Literature Review of Critical Thinking Measurements Used in an Accounting Curriculum

Critical thinking skills have been examined by Kealey et.al. [5] using a writing assignment with a critical thinking rubric to 178 students enrolled in nine sections of a principles of accounting course. The study used seven criteria measuring critical thinking skills as identified by the National Center for Education Statistics (NCES). Specifically, critical thinking skills include interpretation, analysis, evaluation, inference, presenting arguments, reflection, and dispositions. The authors found a significant relationship between critical thinking skills and performance on all of the examinations administered in the course.

In an earlier study by Jenkins [4], critical thinking skills were examined as a predictor of performance for an upper level accounting course. The author used the Watson Glaser Critical Thinking Appraisal (WGCTA), an objective measurement of critical thinking skills. Critical thinking skills were significantly related to performance on two of the three examinations administered in the course.

METHODOLOGY

The objective of this research is to assess the inter-method reliability of critical thinking measurements used in an introductory accounting course. Students in an introductory course in accounting will be given two separate measurements on critical thinking skills. One measurement will consist of a battery

of objective questions using a multiple choice format. A second measurement will use open-ended questions where the students responses will be evaluated using a rubric to assess their critical thinking skills.

Approximately 170 business school students enrolled in the introductory course in Managerial Accounting at an AACSB accredited university will be assessed on their performance in answering questions purported to assess critical thinking skills.

The following hypothesis will be tested with this research:

> H1: There is a strong positive correlation between students' performance on correctly answering multiple choice critical thinking questions and correctly answering open-ended critical thinking questions.

ACKNOWLEDGEMENTS

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

REFERENCES

[1] Arum, Richard and Roksa, Academically adrift, The University of Chicago Press, 2011.

[2] American Institute of Certified Public Accountants (AICPA). "The AICPA Core Competency Framework for Entry into the Accounting Profession", last accessed November 22, 2010 at: http://www.aicpa-eca.org/library/ccf/default.asp?

[3] Baril, C.P., Cunnhgam, B.M., Fordham, D.R., Gardner, R.L. & Wolcott, S.K. (1998). Critical thinking in the public accounting profession: Aptitudes and attitudes. *Journal of Accounting Education*, 16(3/4), 381-406.

[4] Jenkins, E. (1998). The significant role of critical thinking in predicting auditing student's performance. *Journal of Education for Business, 73* (5), 274-280.

[5] Kealey, B. T., Holland, J., & Watson, M. (2005). Preliminary Evidence on the Association between Critical Thinking and Performance in Principles of Accounting. *Issues in Accounting Education, 20* (1), 33-49.

APPENDIX

Sample of Critical Thinking Measurements

Objective questions

1. Select the **incorrect** break-even equation from the following:
A. Total contribution margin = total fixed costs
B. Total contribution margin = total variable costs
C. Total fixed costs/contribution margin ratio = break-even sales in dollars
D. Total revenue = total costs

2. Once sales reach the break-even point, each additional unit sold will:
A. increase fixed cost by a proportionate amount.
B. reduce the margin of safety.
C. increase profit by an amount equal to the per unit contribution margin.
D. increase the company's operating leverage.

3. Zero, Inc. produces a product that has a variable cost of $7.00 per unit. The company's fixed costs are $40,000. The product sells for $12.00 a unit and the company desires to earn a $20,000 profit. What is the volume of sales in units required to achieve the target profit? (Do not round intermediate calculations.)
A. 2,000
B. 8,000
C. 8,500
D. 12,000

4. Which of the following is **not** an assumption made when performing cost-volume-profit analysis?
A. Number of units produced is greater than the number of units sold.
B. Worker efficiency is held constant.
C. The company produces within the relevant range of activity.
D. There is a linear relationship between cost and volume for both fixed and variable cost.

Open Ended Questions

1. Company D has contribution margin of $25 per unit and a break-even point of 10,000 units. If D sells 9,999 units, what would be its net income or loss? Explain how you calculated your answer.

2. What are the assumptions on which cost-volume-profit analysis is based?

3. How does a decrease in variable costs affect the break-even volume of sales?

The Impact of Credit Terms on Supply Chain Cash Flows - A Simulation Approach

Chih-Yang Tsai

School of Business, State University of New York at New Paltz

1 Hawk Drive, New Paltz, NY 12561, USA, tsaic@newpaltz.edu, Tel: +1 845 2572934

Abstract

This study adopts a simulation approach to compare the impacts of a few popular trade-credit terms on the operating cash flow of a manufacturer situated at the middle tier of a three-tier supply chain. The first comparison compares three different credit terms having the same present value from all future cash collections under a stochastic demand process. The simulation results show that despite the equivalence in their present values, the timings and sizes of their cash inflows are different. For small businesses operating on tight cash resources, the timing of cash flow is a critical survival factor. The second comparison adds an additional risk factor, the collection risk, which further delays or reduces the cash inflow and significantly increases the risk, measured by the standard deviation. The simulation approach allows decision makers to project expected future cash flows and risks under various credit terms. It can be used as a tool for crafting fair trade-credit terms between trading parties.

Keywords: Trade credit terms, Cash flows, Supply chain, Simulation

1 Introduction

Cash is one of the most critical resources for business survival. Small businesses, in particular, are constantly battling with the stresses and strains of not having adequate cash resources to grow their businesses during economic boom times and to stay afloat during economic downturns. Due to the trade credit terms offered in common business practices, the timing of cash flow does not coincide with that of sales transactions, making it harder for managers to forecast cash flows especially with the risks of late collection and default. Small businesses, especially those dealing with larger customers and suppliers, are often squeezed by both sides, i.e., customers asking for extended credit terms and suppliers demanding shorter credit periods. The two factors, together with tightened credit lines by their banks, become a lethal combination to many small businesses [15]. Recommendations on ways to improve cash flow [12] following the concept of reducing the cash conversion cycle are often difficult for small businesses to implement as they lack the bargaining power to negotiate for favorable credit terms [18].

To highlight the impact of trade credit terms on cash inflow, this study first compares three credit terms that are equivalent from the net present value point of view for their future cash collections of an invoice. We assume a trading environment where the manufacturer makes production decisions based on its forecast of future demand. Customers place orders and receive delivery in the same time period and all invoices are issued at delivery. As such, forecasting errors may incur inventory cost for overproduction and backlogged cost for shortages. Demand forecast and cash collection/disbursement are aggregated from all products and invoices on the total dollar value for each decision period, say, a month. This is analogous to the balance-forward method for managing accounts receivable (AR).

The three basic types of credit terms considered in this study include

1. net terms: Payment is due r periods after delivery or invoice date.

2. cash-on-delivery (COD) with a discount: Payment is due on delivery with a discount rate c_r. This can be considered as factoring where the factors reimburse the manufacturer on the value of AR minus a fixed percentage charge, c_r, while taking over the ownership of the accounts receivable.

3. cash-before-delivery (CBD) plus net terms: The manufacturer charges a fixed percentage, w, of customers' future order amount as a down payment at the time production starts. In exchange for the down payment, the manufacturer offers an extended credit period, $r + x$, where x is the additional credit period for the balance.

The CBD plus net terms gained its popularity during economic downturns. The Enterprise Recourses Planning (ERP) system by SAP provides the functionality of processing such credit terms as stated on its web site, "Often in business, especially in a make-to-order environment, customers may be required to pay some amount in advance before delivery of goods." This ERP process allows the seller to create requests for the down payment, to record the down payment, and to generate a final invoice for the balance [14].

In the first comparison, we find the break-event discount and down payment rates, $\bar{c}_r$ and $\bar{w}$, such that the three types of credit terms produce the same present value for their future cash collections, given r, x, and l where l is the production lead time. In the simulation study, we adopt a stochastic demand pattern with a positive trend because cash resources are the tightest during a growth period [17].

In addition to the aforementioned three types of credit terms, we also consider a two-part term where the customer is offered a discount to encourage early payment in the second comparison. The most common two-part terms are (1/10, net 30) and (2/10, net 30) [13]. In the first comparison, we do not include this option because the portion of early payment is stochastic. In the second comparison, we add the collection risk and assume that the manufacturer can estimate the proportions of open invoices collected on time, late, or defaulted from historical data. We compare two types of two-part terms. The first is a traditional two-part terms with an early payment discount of c_r and net on or before r periods. The second type has a CBD down payment due when production starts and the balance of the order amount is invoiced under a two-part term but with an extended credit period of $r + x$. The remainder of this paper is organized as follows. Section 2 provides a

brief literature review, followed by Section 3 where the basic model is introduced. Section 4 and Section 5 conduct comparison 1 and comparison 2 respectively. Section 6 summarizes the findings of this study.

2 Literature Review

Mian and Smith [9] examined the factors that influence a company's accounts receivable policies from the reasons to extend trade-credit to the choices of AR financing sources and using captives. An empirical study [13] surveyed the trade credit practices in three countries, US, UK, and Australia, found that the most popular trade credit period is 30 days. The study also verified the two main reasons for offering trade credit terms, namely, as a buffer to treat the information asymmetry between buyers and sellers, and as an instrument for price discrimination. The authors also found a higher percentage of US companies offering the two-part credit terms than those in the other two countries. For extensive discussions of the rationales of offering credit terms and general practices, readers are encouraged to read [11], [18], and [13].

Several studies showed that cash flows, especially operating cash flow is a better indicator of a company's solvency. Mills and Yamamura [10] argued that auditors should shift some of their attentions from balance sheet and income statement to cash flow statement in order to assess a company's liquidity and recommended the use of a few cash flow ratios. Almeida, *et al.* [1] proved empirically that the cash flow sensitivity of cash (the propensity to save cash out of cash inflow) is a better proxy to measure a company's financial constraints. A financially constrained company tends to implement a financial policy that systematically holds cash for future needs, while an unconstrained company does not show such pattern. For companies that are not financially sound, analysts often produced operating cash flow forecasts as additional information to assist security valuation [5]. Dechow, *et al.* [4] showed that current earning is a better predictor for future operating cash flow than current operating cash flow, using both analytical and empirical approaches.

In supply chain literature, a first-order autoregressive, $AR(1)$, demand pattern has been used in several studies [7, 8, 3, 16, 2]. In a study showing the cash strains during a growth period, Tsai [17] adopted a modified $AR(1)$ demand pattern with a trend to capture a growth factor. The linkage between cash management and supply chain management is best seen through the cash conversion cycle [6].

3 The Model

A first-order autoregressive model, $AR(1)$, with trend is adopted for this study [17] in which the demand of time period t in dollar amount is determined by Equation 1 below.

$$D_t = \beta_0 + \beta_1(t) + \rho D_{t-1} + \epsilon_t, \quad t = 1, 2, \ldots \tag{1}$$

where β_0 and β_1 are the intercept and slope of a linear trend respectively and $\rho D_{t-1} + \epsilon_t$ is a first order autoregressive process. The manufacturer adopts a produce-up-to policy where the produce-

up-to quantity is the expected lead time demand plus a safety stock [8]. Under this assumption, the manufacturer determines the production volume, Y_t, to be started in time $t+1$ by

$$\begin{aligned} Y_t &= D_t + (S_t - S_{t-1}) \\ &= D_t + (E[\sum_{k=1}^{l+1} D_{t+k}|D_t] - E[\sum_{k=1}^{l+1} D_{t-1+k}|D_{t-1}]) \end{aligned} \quad (2)$$

where S_t and S_{t-1} are the produce-up-to quantities estimated in time t and $t-1$ respectively. Production of Y_t is completed in $t+l$ where l is the production lead time. Thus, Y_t is available for sale in $t+l+1$.

We assume the unit cost of materials is a fixed percentage, α_1, of unit sale price and the manufacturer always pays its suppliers at the end of the credit period, v. There is no purchase lead time. However, all materials need to arrive before production starts. For example, materials for producing Y_t need to arrive in period t for the production to begin in $t+1$. We also assume unit internal operations cost, which includes labor cost and other overheads allotted to each unit, is a fixed percentage, α_2, of unit sale price and the payment of this cost is split evenly across the periods in the lead time. In addition, holding cost and backlog rates per period, h and s, are fixed percentages of the sales value of the units held and backlogged respectively.

The following Theorem contains several key results borrowed from [17] for this study.

Lemma 3.1. *According to the demand process in Equation (1), expected future demand D_{t+i} and its associated variance can be expressed in terms of D_t once D_t is realized .*

$$E[D_{t+i}|D_t] = \rho^i D_t + \frac{1-\rho^i}{1-\rho}[\beta_0 + \beta_1(t+i)] - \beta_1 \sum_{k=1}^{i-1} k\rho^k \quad (3)$$

$$Var[D_{t+i}|D_t] = \frac{\sigma^2}{1-\rho^2}(1-\rho^{2i}) \quad (4)$$

The expected production amount for a future period $t+i$ estimated at time t can be written as:

$$E[Y_{t+i}] = \frac{1-\rho^i}{1-\rho}\beta_0 + \sum_{k=1}^{i} \rho^{i-k}\beta_1(t+k) + \rho^i Y_t + \frac{1-\rho^i}{1-\rho}\beta_1(l+1) \quad (5)$$

When determining the amount of down payment at time period t, we use wY_t instead of $wE[D_{t+l+1}|D_t]$ for two reasons.

1. $E[D_{t+l+1}|D_t]$ is the forecast for future demand in time period $t+l+1$ based on the $AR(1)$ process. With a growth trend, $E[D_{t+l+1}|D_t] > D_t$. It would be difficult to convince the buyers to pay a down payment based on a higher demand level which may or may not happen l periods later.

2. On the other hand, Y_t is based on the actual production level, which is visible right now. And it is based on the current level of demand with an adjustment of the produce-up-to amount from the previous period (see 3). Thus, it is easier for customers to comprehend.

4 Comparison I: break-even terms

Given the key elements in trade credit terms, regular credit period r, extended period $r + x$, and lead time l, let $\bar{w}$ and $\bar{c}_r$ be the break-even w and c_r such that the present values of all the three types of credit terms mentioed earlier are the same. For any order size D, the break-even between the net terms and CBD plus net terms implies

$$\bar{w}D + (1 - \bar{w})DK^{l+r+x} = DK^{l+r}$$

where $K = (1 + k)^{-1}$ is the discount rate per period and k is the capital cost rate per period.

And, the break-even between the COD with discount terms and net terms implies

$$DK^{l+r} = (1 - c_r)DK^l$$

Hence

$$\bar{w} = \frac{K^{l+r} - K^{l+r+x}}{1 - K^{l+r+x}} \quad (6)$$

$$\bar{c}_r = 1 - K^r \quad (7)$$

From (6) and (7), we learned that

- $\bar{w}$ decreases as l increases: When the lead time is long, the down payment provides added benefits to the manufacturer to counter the long days in inventory. Thus, the break-even $\bar{w}$ decreases.
- $\bar{w}$ decreases as r increases: When the original credit term r is long, the down payment provides added benefits to the manufacturer to counter the long days in receivables. Thus, the break-even $\bar{w}$ decreases.
- $\bar{w}$ decreases as x decreases: When the extension of the credit period x decreases, it shortens the days in receivables. Thus, $\bar{w}$ decreases.
- $\bar{c}_r$ increases as r increases: When the credit period, r, is longer, it is beneficial for the manufacturer to raise the discount rate, c_r, to incentivize early payment.

Assume all customers pay the invoiced amount at the set times according to the credit terms, the only cash flow risk comes from demand uncertainty. All ARs are eventually fully collected one way or another.

4.1 Simulation Results

We use the following parameters as the baseline pattern to illustrate the impact of credit terms on cash flows.

1. demand parameters: $\beta_0 = 100$, $\beta_1 = 5$, $\rho = 0.6$, $\sigma = 5$; The historical demand $D_1, \ldots, D_{30}$, i.e. $t = 30$, are randomly generated according to Equation (1).

2. time related parameters: production lead time $l = 2$; AR credit period $r = 2$; AP credit period $v = 2$; extra credit period with down payment $x = 2$.

3. cash outflow parameters: material cost rate, $\alpha_1 = 0.3$; internal processing cost rate, $\alpha_2 = 0.6$; annual capital cost rate, 8%, i.e., $k = \frac{8\%}{12}$; holding cost rate per period, $h = 1\%$.

4. cash inflow parameters: Backlog discount rate per period, $s = 10\%$.

Based on the given parameters, $\bar{c}_r = 1.32\%$ and $\bar{w} = 32.9\%$ from (6) and (7). In addition to the three break-even scenarios, we add the fourth scenario, which is the same as scenario two except that there is a deeper discount. The four scenarios are summarized as follows.

1. Scenario 1 (break-even net terms): all collections made 2 periods (r) after delivery,

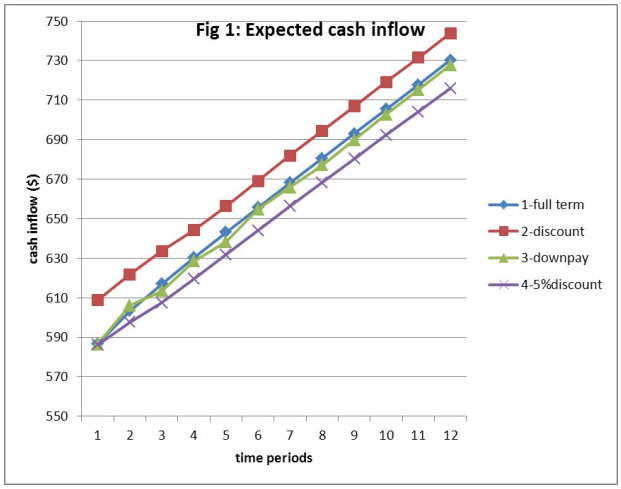
2. Scenario 2 (break-even cash-on-delivery with discount):all collections made at delivery with a 1.32% discount,

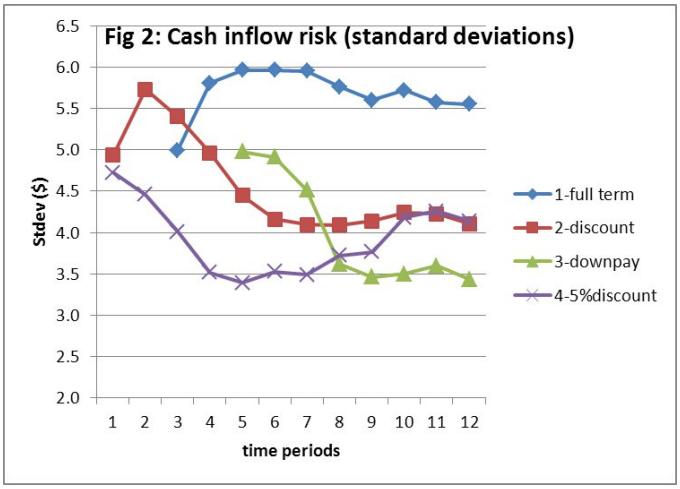
3. Scenario 3 (break-even cash-before-delivery and net terms): 32.9% of order amount paid 2 periods (l) before delivery when the production starts and the balance paid 4 periods, $r + x$, after delivery,

4. Scenario 4 (deep discount for cash-on-delivery): the same as scenario 2 but with a deeper discount $c_r = 5\%$.

Since the credit period for AP is the same as the lead time, the suppliers fully finance the manufacturer for its material cost. Because actual demand is a random variable, all variables derived from it are random variables too. We use the simulation process to estimate key statistics of those random variables. The simulation process runs 1000 iterations to estimate the means and standard deviations of actual demand and cash flows in a 12-periods horizon from period 31 to period 42. As the outflow parameters are the same for all scenarios, we focus on cash inflow in this analysis. Fig 1 shows the expected cash inflow in the 12-period planning horizon. It is clear that Scenario 2, having the fastest collection of AR, produces the highest cash inflow. Scenarios 1 and 3 are very close to each other. While Scenario 1 has a shorter credit period, Scenario 3 collects a portion of sales amount, $\bar{w}$, before delivery. Scenario 4 produces the worst result as the discount rate is very high, indicating that if the cost of factoring is too high, it might hurt the cash inflow despite a fast recovery of AR.

Fig 2 depicts the cash inflow risks among the four scenarios. Cash flow in the first few periods exhibits some randomness due to the impact from realized demands, $D_{30}, D_{29}, \ldots$ in the past. The longer the credit period, the longer the impact as the manufacturer is collecting mostly from realized demands in the first few periods. The standard deviations in later periods show that Scenario 1 has the highest risk. In general, the shorter the credit period, the lower cash inflow risk because demands further into the future have a higher risk (see Theorem 4). However, Scenario 3 has a lower standard deviation because 32.9% (w) of the order value has been collected through the down payment and does not involve any risk at all.





This comparison shows that credit terms which seem equivalent could produce very different cash inflows. Scenario 2 which has the lowest days in receivables produces the best overall result, highest cash inflow and moderate risk.

5 Comparison II: with collection risk

In the previous analysis, we assume collections are certain in their amount and timings. This section considers the collection risk in addition to the demand risk. The collection risk includes the risks of early collection, on time collection, late collection and default. The focus here is on two-part credit terms. First part is a discount incentive for early payment and if the customer decide not to take that option, the second part is a typical net term. It is assumed that there is no incentive for customers to pay before the full term once they missed the discount cut-off time. The two-part terms have an added uncertainty contributed by not knowing the proportion of early collection.

Tsai [16] demonstrated that although early collection improves cash inflow, it also increases the risk due to the added uncertainty.

Let p_0 be a random variable representing the proportion of early collection (within the discount period); p_1 represents the proportion of the remaining balance collected at full term or first collection attempt, and p_i, $i = 2, 3, \ldots$ denote the proportion of the uncollected balance collected in time $r + (i-1)$ period after the invoice date which is $i-1$ periods overdue upon the i^{th} collection attempt. We assume all p_i, $i = 0, 1, \ldots$ are independent normally distributed random variables with mean μ_i and variance σ_i^2 and the manufacturer is able to estimate μ_i and σ_i^2 from previous experience. After certain number of collection attempts, say N, $p_i = 0$ for $i > N$, either because there is no balance left or because the balance is deemed defaulted.

5.1 Simulation results

We use the same demand pattern described in the previous section for this simulation on two scenarios. Scenario A is a pure two-part term and Scenario B is a cash-before-delivery plus two-part term. Since we introduce the collection risk in this analysis, the setups of Scenario 2 and Scenario 4 in the previous section are no longer appropriate because they do not have credit periods. Instead of using $\bar{w}$ and $\bar{c}_r$, we adopt the popular values in practice where $w = 10\%$ and $c_r = 2\%$. Collection parameters for the two scenarios are:

- Scenario A: $r = 2$; $\mu_0 = 0.2$, $\mu_1 = 0.8$, $\mu_2 = 0.6$, $\mu_3 = 0.1$, $\mu_i = 0$ for $i > 3$
- Scenario B: $r = x = 2$; $\mu_0 = 0.1$, $\mu_1 = 0.8$, $\mu_2 = 0.6$, $\mu_3 = 0.2$, $\mu_i = 0$ for $i > 3$
- For both scenarios, we assume $\sigma_i = 0.25\mu_i$ for $i = 0, 1, 2, 3, \ldots$.

Note that the values of p_0 and p_3 are switched between the two scenarios since the requirement of a down payment in Scenario B may lower the motivation for early payment.

Fig 3 indicates that except for the few early periods heavily influenced by the realized past demands, Scenario A has a higher expected cash inflow as it has a shorter collection cycle, which cannot be countered by the 10% down payment in Scenario B. In addition, the amount of inflow in each period is lower than that of the same period in all three scenarios of comparison 1 due to the delay and default factors of the collection risk. Consistent with the previous comparison, Fig 4 indicates that Scenario B has a lower risk as a result of the risk free down payment. Because of the added collection risk, the magnitude of the inflow risk in this comparison is much higher than that observed in Fig 2.

Although this study focuses on the accounts receivable, the simulation design does include the cash outflow factors, such as the payment to suppliers, internal operations cost, inventory cost and backlogging cost. Because we assume the manufacturer always pays its suppliers at the full term of its AP and internal cost is spread evenly across the entire lead time periods, all cash outflow risks are derived from the demand risk. The two lines in the middle of Fig 5 represent the expected cash balance of the two scenarios. Cash balance is the cumulative cash inflow minus cumulative cash outflow. Here, we assume the beginning cash balance of the two scenarios is the same. The two lines above the expected cash balance lines represent the expected balance plus two standard

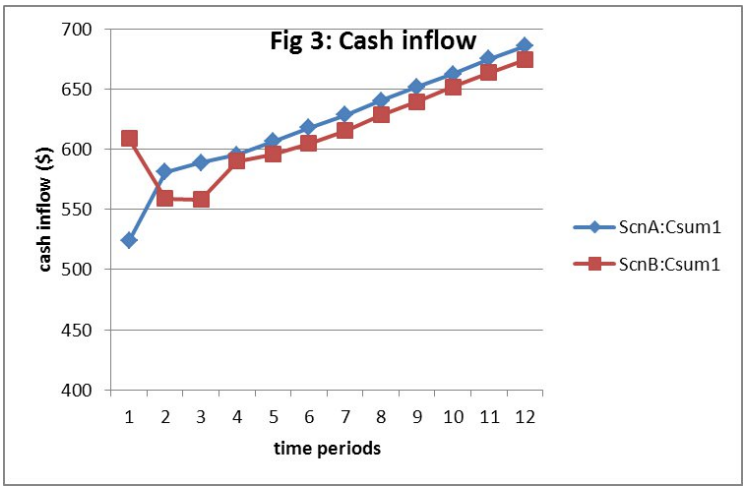


Fig 3: Cash inflow
700
650
600
550
500
450
400
cash inflow ($)
1 2 3 4 5 6 7 8 9 10 11 12
time periods
ScnA:Csum1
ScnB:Csum1

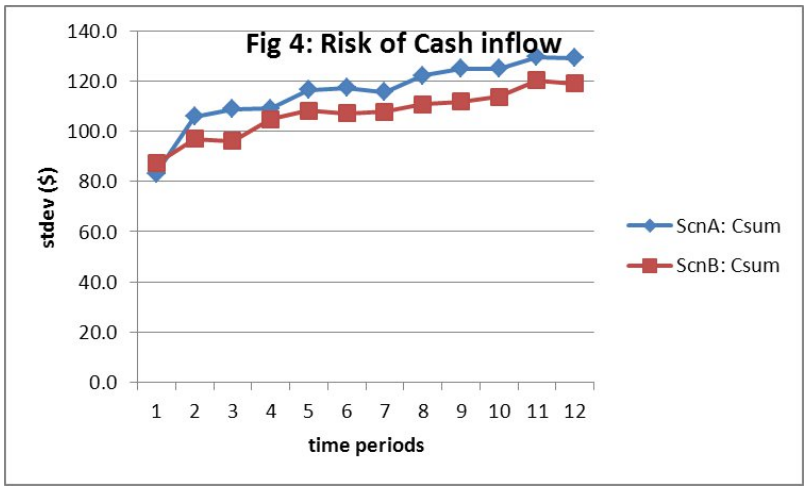
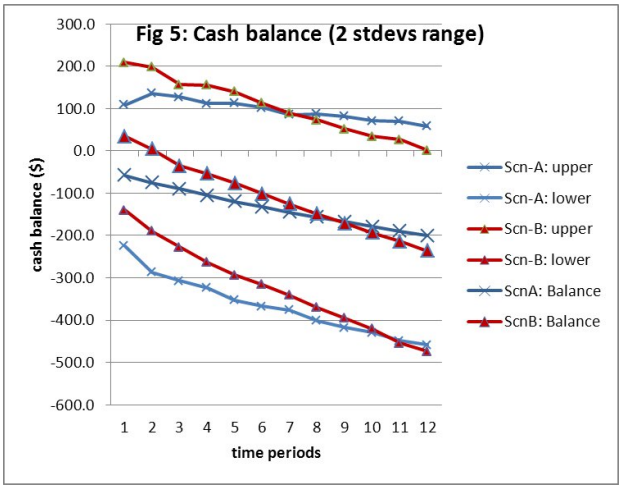


Fig 4: Risk of Cash inflow
140.0
120.0
100.0
80.0
60.0
40.0
20.0
0.0
stdev ($)
1 2 3 4 5 6 7 8 9 10 11 12
time periods
ScnA: Csum
ScnB: Csum



deviations and the two lines below show the expected balance minus two standard deviations. When looking at the later periods in the planning horizon, Scenario B has a narrower range in terms of cash balance due to the risk free down payment.

Fig 5 also indicates that during a growth period, the rise of cash outflow out-paces the increase of inflow due to the credit period, resulting in a declining cash balance. It also suggests that when customers push for a longer credit period by offering a small down payment, the manufacturer suffers from a lower expected cash inflow but gains on the lower risk. Sometimes, lowering risk level is as good as increasing cash balance. When making cash flow related decisions, one must consider both expected flows and risks. A lower flow with a smaller risk might work better than a higher flow with a larger risk in some cases. The former allows a company to set a lower cash balance service level. For example, a company implements a policy of maintaining a minimum cash reserve so that the chance of its cash balance falling below the reserve level is less than 5%. Under such circumstances, a cash flow distribution with high risk requires a higher level of cash reserve, which further limits the utilization of the scarce cash resources.

6 Conclusion

This study provides two comparisons among different credit terms under different risk assumptions. The first comparison shows that when considering trade credit terms, better timing of cash inflow is more important than the present value of future flow. The second comparison introduces collection risk into credit terms decisions. It indicates that when risk is a significant factor, a sound credit term decision must balance between expected flows and risks.

The simulation approach implemented in this study allows decision makers to assess expected cash flows and their associated risks under various credit terms. As a result, it can be used as a decision tool when negotiating trade credit terms with customers or suppliers. A decision maker can compare the results from different scenarios to help identify reasonable terms fair and acceptable by all trading parties.

7 References

References

[1] H. Almeida, M. Campello, and M. S. Weisbach. The cash flow sensitivity of cash. *Journal of Finance*, 59(4):1777–1804, 2004.

[2] S. Balan, P. Vrat, and P. Kumar. Information distortion in a supply chain and its mitigation using soft computing approach. *Omega: The International Journal of Management Science*, 37(2):282–299, 2009.

[3] F. Chen, Z. Drezner, J.K. Ryan, and D. Simchi-Levi. Quantifying the bullwhip effect in a simple supply chain: the impact of forecasting, lead times, and information. *Management Science*, 46(3):436–443, 2000.

[4] P.M. Dechow, S.P. Kothari, and R.L. Watts. The relation between earnings and cash flows. *Journal of Accounting & Economics*, 25:133–168, 1998.

[5] M.L. DeFond and M. Hung. An empirical analysis of analysts' cash flow forecasts. *Journal of Accounting & Economics*, 35:73–100, 2003.

[6] M.T. Farris and P.D. Hutchison. Cash-to-cash: the new supply chain management metric. *International Journal of Physical Distribution & Logistics Management*, 32(4):288–298, May 2002.

[7] H.L. Lee, V. Padmanabhan, and S. Whang. Information distortion in a supply chain: The bullwhip effect. *Management Science*, 43(4):546–558, April 1997.

[8] H.L. Lee, K.C. So, and C. Tang. The value of information sharing in a two-level supply chain. *Management Science*, 46(5):626–643, 2000.

[9] S. L. Mian and C. W. Smith. Accounts receivable management policy: Theory and evidence. *Journal of Finance*, 47(1):169–200, 1992.

[10] J. R. Mill and J. H. Yamamura. The power of cash flow ratios. *Journal of Accountancy*, 186(4):53–61, 1998.

[11] C. K. Ng, J. K. Smith, and I. L. Smith. Evidence on the determinants of credit terms used in interfirm trade. *Journal of Finance*, 54(3):1109–1129, 1999.

[12] NYSSCPA. 10 ways to improve small business cash flow. *Journal of Accountancy*, page 14, March 2000.

[13] R. Pike, N.-S. Cheng, K. Cravens, and D. Lamminmaki. Trade credit terms: Asymmetric information and price discrimination evidence from three continents. *Journal of Business Finance & Accounting*, 32(5 & 6):1197–1236, 2005.

[14] SAP. Sales order processing with customer down payment. http://help.sap.com/bp_dm603/BBLibrary/HTML/201_EN_US.htm, 2011. retrieved on Oct. 4, 2011.

[15] John Tozzi. A cash-flow crisis is the recessionâĂŹs legacy. *BusinessWeek*, March 28- April 3:59–60, 2011.

[16] C-Y Tsai. On supply chain cash flow risks. *Decision Support Systems*, 44:1031–1042, 2008.

[17] C-Y Tsai. On delineating supply chain cash flow under collection risk. *Int. J. Production Economics*, 129:186âĂŞ194, 2011.

[18] N. Wilson and B. Summers. Trade credit terms offered by small firms: Survey evidence and empirical analysis. *Journal of Business Finance & Accounting*, 29(3 & 4):317–351, 2002.

GETTING OFF ON THE RIGHT FOOT: THE IMPORTANCE OF EARLY SUCCESS FOR ACCOUNTING MAJORS

William E. Bealing, Jr., Bloomsburg University, 570-389-4386
Richard L. Baker, Bloomsburg University (Emeritus)

ABSTRACT

This study examines the relationship between an accounting major's performance, as measured by course grade, in their introductory accounting course and performance in subsequent courses. A positive relationship was postulated and found to exist. This relationship held even after controlling for a student's "quality", as measured by SAT scores. This relationship applied to both quantitative and qualitative courses. Additional analysis revealed that initial performance in the introductory accounting course was a significant predictor of whether a student would graduate from the program within four years, as well as predicting a student's GPA. Two theories are presented to explain these findings.

INTRODUCTION

The students' innate abilities and motivation are often cited as the key elements to determining their success in an undergraduate accounting program. However, these traits are hard to measure because it is often impossible to determine what is going on inside a student's head. Prior research has examined the link between personality type and performance in a variety of accounting courses [3] [9] [10] [11] [12]. Specifically, one particular personality type, sensing-judging (SJ) is related to successful performance in a variety of courses [3]. They hypothesized that personality type could be used as a surrogate measure of unquantifiable characteristics, such as motivation and ability. However, could there be other readily available indicators that would accurately predict a student's performance in a given course?

This paper extends earlier lines of research by examining the theory that grades received in a student's first accounting course can predict performance in subsequent accounting and non-accounting courses. When students begin their first accounting course, presumably, the best available indicators of their potential performance are their SAT scores. However, once students complete their first accounting course, their grades provide relevant new information which can be used to predict their future performance. In the spirit of Bayesian statistics, this new information is used to "update" the assessment of a student's performance in subsequent accounting and other business core courses. Prior research has found that grades earned in prerequisite accounting courses were statistically significant linear predictors of performance in subsequent accounting courses [3]. This paper extends this line of research by positing that a student's performance in their introductory accounting course should be positively correlated to performance in all required subsequent accounting and non-accounting business courses.

In the area of accounting, prior research has found that students' grade point averages (GPA), performance in their accounting principles course, and performance in their first statistics course, were all significantly correlated to their success in cost accounting [7].

It is interesting to note that performance in prior related coursework does not predict performance in college level information systems (IS) courses [8]. This research looked at students who had IS courses in high school and correlated it to later performance in an introductory college level IS course. These results are consistent with one of the first studies to specifically look at high school accounting courses and subsequent performance in the first college accounting course. In that study, no difference was found in the performance of students who had bookkeeping courses in high school versus those who did not [2]. However, other researchers have found that students' pre-college study of accounting/bookkeeping was significantly related to their performance in an introductory accounting course in college [6].

Other researchers have also found that student performance in certain courses was associated with performance in specific courses taken subsequently. For instance, students who had taken calculus and related mathematics courses performed better in economics and finance courses than individuals who had no prior coursework in the field of mathematics [1] [4]. While other studies concluded that completed accounting courses were significantly correlated with student performance in other business courses [5].

This logic is intuitively appealing and has implications for the accounting curricula at a typical university. Most accounting professors purport to require students to be able to think critically, rather than perform rote memorization. If this is true, performance in an introductory accounting course should provide insight into student performance in subsequent accounting courses. But, could this information also predict performance in non-accounting, quantitative courses, such as economics and statistics? Furthermore, could student performance in their initial accounting course also be used to explain their subsequent performance in courses in such diverse fields as Management, Business Law, and even Public Speaking?

METHODOLOGY

This study is based on data collected from incoming freshman Accounting majors over the last seven years at one, mid-sized, comprehensive, public university. The research for this paper began with a review of the accounting department's list of courses required for graduation. Nineteen "common" courses were identified and used as the basis for testing of:

H1 There is no significant positive correlation between performance, as measured by course grade, in Principles of Accounting I and performance, as measured by course grade, in eighteen other required courses.

The nineteen courses which form the basis of this study are identified in Exhibit 1.

The letter grades received by students on their first attempt at each of the nineteen courses were numerically coded using the following scale: A=4; B=3; C=2; D=1; E=0. Grades of W (withdrew), I (incomplete) and TR (grade transferred in) were not coded in the initial analysis. Subsequently, all of the analyses were repeated with grades of W and I being assigned a value of 0. There were no statistical differences between the two sets of analyses.

Exhibit 1
Courses Required of Accounting Majors

#	Required Course	Accounting/Business	Q /NQ*
1	Principles Accounting I	Accounting	Q
2	Principles Accounting II	Accounting	Q
3	Managerial Accounting	Accounting	Q
4	Cost Accounting	Accounting	Q
5	Intermediate Accounting I	Accounting	Q
6	Intermediate Accounting II	Accounting	Q
7	Intermediate Accounting III	Accounting	Q
8	Tax	Accounting	Q
9	Auditing	Accounting	Q
10	Economics I	Business	Q
11	Economics II	Business	Q
12	Bus and Econ Math	Business	Q
13	Bus and Econ Stat	Business	Q
14	Matrix Algebra	Business	Q
15	Business Law	Business	NQ
16	Principles of Management	Business	NQ
17	Principles of Marketing	Business	NQ
18	Intro to Corp. Finance	Business	Q
19	Public Speaking	Business	NQ

***Q=Quantitative NQ=Non-Quantitative**

Although the university employs a plus/minus grading system the determination was made to disregard any plus or minus in the grade. For example, if a grade of C+ was earned, it was treated as a "C" for purposes of this study. This was done because while a plus/minus system is currently in place, the transition to the system took place after the study was begun and the decision was made to continue coding results as if the five letter grading system had continued.

Both simple and multiple regression models were employed to examine the relationship between the individual students performances in their first accounting course (Principles of Accounting I) and the remaining eighteen required courses.

RESULTS

Table 1 contains the results of the analysis of the eight required accounting courses which are taken by accounting majors, subsequent to Principles of Accounting I (Principles I). All the regression analyses yielded significant results in the anticipated direction. Closer examination of Table 1 (Column A) reveals adjusted R-Square (R2) values ranging between approximately .05 and .30 where the students' grades in Principles I were the independent variable. Thus, by inference, the grade a student received in their Principles of Accounting I class explained between 5% and 30% of the variation of their grade received in specific subsequent accounting classes.

TABLE 1
Principles of Accounting I as a Predictor of Success in Subsequent Accounting Courses

Course	A Principles I	B SAT Score	C SAT/Principles I Both
Principles Acctg II	**p=<.0001* R2=.3040**	**p=.0002* R2=.0493**	**Prin I p=<.0001* SAT p=.1605**
Managerial Acctg	**p=<.0001* R2=.1787**	**p=<.0001* R2=.0739**	**Prin I p=<.0001* SAT p=.0003* R2=.2251**
Cost Accounting	**p=.0007* R2=.0913**	**p=.0050* R2=.0784**	**Prin I p=.0603*** **SAT p=.0380**** **R2=.1334**
Intermediate I	**p=<.0001* R2=.1811**	**p=.0404* R2=.0215**	**Prin I p=.0001* SAT p=.7449**
Intermediate II	**p=<.0001* R2=.1415**	**p=<.0001* R2=.1406**	**Prin I p=.0023* SAT p=.0050* R2=.1847**
Intermediate III	**p=.0128**** **R2=.0505**	**p=.0541*** **R2=.0329**	**Prin I p=.0605*** **SAT p=.0891*** **R2=.0934**
Federal Taxation	**p=.0018* R2=.0593**	**p=.0004* R2=.0909**	**Prin I p=.0636*** **SAT p=.0186**** **R2=.0997**
Auditing	**p=.0062* R2=.0835**	**p=.0052* R2=.1217**	**Prin I p=.1350 SAT p=.0601***

***S=Signif at p<.01; **S=Signif at p<.05; ***S=Signif at p<.10; R2=R-squared (Adj)**

However, one could make the argument that "a student's innate ability will determine his/her performance in a given course". For decades, total SAT score has been used as a quantifiable determinate of innate ability. Therefore, the regression models were re-coded to examine the relationship between a student's SAT score and their performance in the same eight accounting courses. Table 1 (Column B) contains the results of the regression models which specified student performance in the eight required accounting courses are a function of SAT score. Once again, all of the regression analyses yielded significant results in the anticipated direction. At this point, in order to account for the impact of "innate ability", multiple regression models were run to examine students' performance in the eight required accounting courses as a function of both their performance in Principles I and their SAT score. These results are presented in Table 1 (Column C).

Both the Principles I and SAT scores proved to be significant in explaining performance for Managerial Accounting, Cost Accounting, Intermediate II, Intermediate III and Federal Taxation. However SAT score was not significant, while Principles I performance still emerged as significant in explaining performance for Principles II and Intermediate I. This means that even when a student's SAT score is taken into account, Principles I remains a significant predictor of performance in seven of the eight required Accounting courses. Only in the case of Auditing,

was Principles I not significant (at p<.10 levels) when a student's SAT scores were also considered.

Table 2 contains the results of the analysis of the six quantitative, non-accounting courses all accounting majors are required to complete. Column A reports the results when Principles I grades were run as the independent variable while the results for the SAT scores are reported in column B. Finally, column C contains the results of multiple regression utilizing both Principles I grades and SAT scores as independent variables.

All of the simple regression models yielded statistically significant results in the anticipated direction for Principles I with adjusted R-Square (R2) values ranging between approximately .08 and .31. Thus, by inference, the grade a student received in their Principles of Accounting I class explained between 8% and 31% of the variation in their grades received in their quantitative, non-accounting courses. Multiple regression results confirmed that Principles I grades were significant even in the presence of the SAT scores. Once again, the results indicate that a student's performance in Principles of Accounting I yields significant predictive power about how students will perform in their quantitative, non-accounting courses.

TABLE 2
Prin of Accounting I as a Predictor of Success in Non-Accounting, Quantitative Courses

Course	A Principles I	B SAT Score	C SAT/Principles I Both
Principles of Econ I	**p=<.0001* R2=.2027**	**p=<.0001* R2=.0419**	**Prin I p=<.0001* SAT p= .0749*** R2=.2154**
Prin of Econ II	**p=<.0001* R2=.1909**	**p=.0101** R2=.0167**	**Prin I p=<.0001* SAT p= .2590 R2=.1697**
Bus & Econ Math	**p=<.0001* R2=.0848**	**p=.0687*** R2=.0157**	**Prin I p=.0302** SAT p=.1986 R2=.0477**
Bus & Econ Stat	**p=<.0001* R2=.1451**	**p=.0687*** R2=.0157**	**Prin I p=<.0001* SAT p= .6539 R2=.1632**
Matrix Algebra	**p=<.0001* R2=.3051**	**p=<.0001* R2=.0865**	**Prin I p=<.0001* SAT p= .0052* R2=.3428**
Intro Corp Finance	**p=.0004* R2=.0974**	**p=.0522*** R2=.0242**	**Prin I p=.0366** SAT p=.1103 R2=.0719**

***S=Signif at p<.01; **S=Signif at p<.05; ***S=Signif at p<.10; R2=R-squared (Adj)**

Table 3 contains the results of the analyses for the four qualitative business courses required of all accounting majors. Column A shows the results of using Principles I grades as the

independent variable. Column B repeats the analysis, but this time SAT scores are the independent variable. The results indicate that performance in Principles I is a significant predictor for each of the four required qualitative courses with the adjusted R-square (R2) values ranging from approximately .12 to .19.

When the results are examined for the SAT scores (Column B), significant results are found only for Business Law and Principles of Management. Multiple regressions using both Principles I grades and SAT scores were conducted for these two courses. The results, contained in Column C, indicate that Principles I grades were still significant even in the presence of the SAT scores. However, in neither instance were SAT scores a significant variable. No multiple regression models were run for either Public Speaking or Principles of Marketing since the SAT scores were not significant in the simple regression models. The results indicate that grades earned in Principles of Accounting I are a significant predictor of performance in the non-quantitative, non-accounting courses.

TABLE 3
Prin of Accounting I as a Predictor of Success in Non-Acctg, Non-Quantitative Courses

Course	A Principles I	B SAT Score	C SAT/Principles I Both
Public Speaking	**p=<.0001* R2=.1343**	**p=.2849 R2=.0007**	**N/A**
Business Law	**p=<.0001* R2=.1852**	**p=.0008* R2=.0632**	**Prin I p=<.0001* SAT p=.1493**
Prin of Management	**p=<.0001* R2=.1228**	**p=>0616*** **R2=.0144**	**Prin I p=<.0001* SAT p=.4681**
Prin of Marketing	**p=<.0001* R2=.1432**	**p=.1513 R2=.0078**	**N/A**

***S=Signif at p<.01; **S=Signif at p<.05; ***S=Signif at p<.10; R2=R-squared (Adj)**

It was decided to extend the analysis to the examination of the relationship between performance in a student's first accounting course and the ultimate measure of academic success for Accounting majors, graduation with a degree in Accounting. Since many schools use five year graduation rates as a measure of success, a subset of the data consisting of accounting majors who enrolled at the university at least five years ago was used for this analysis. It was assumed that anyone who entered the university at least five years ago should have had an adequate opportunity to graduate from the program. An indicator variable was created which was coded as a 1, if the student successfully completed all their courses on the first attempt within the five year window. The indicator variable was coded as 0, if either the student had not successfully completed all of the courses on their first attempt or had not graduated within the five year window.

Table 4 contains the results of the analysis for successful graduation within five years. The results for simple regression analysis utilizing Principles I grades as the independent variable are reported in Column A and SAT score results are in Column B. Finally, Column C contains the multiple regression results using both Principles I and SAT grades as independent variables. The

simple regression analysis of Principles I grades yielded a significant result in the anticipated direction with an adjusted R-Square (R2) value of .3979. Thus, by inference, the grade a student received in their Principles of Accounting I class explained about 40% of the probability that a student would successfully graduate within five years. While a similar analysis using SAT scores also yielded significant results, the multiple regression analysis (Column C) indicated that Principles I grades were significant but SAT scores were not. Thus, it appears that performance in the Principles I accounting class can provide insight into whether a student will successfully graduate within five years.

TABLE 4
Principles of Accounting I as a Predictor of Five Year Graduation Success

Principles I	SAT Score	SAT/Principles I Both
P=<.0001*	p=.0038*	Prin I p=<.0001*
R2=.3979	R2=.0697	SAT score p=.1016

***S=Signif at p<.01; **S=Signif at p<.05; ***S=Signif at p<.10; R2=R-squared (Adj)**

But could a student's performance in the initial accounting course also tell us something about his/her potential grade point average (GPA) in the other eighteen business courses under review in this research project? At this point in the research study, two separate analyses were conducted. The first looked at the students' GPAs for the eighteen courses as a function of both their Principles I grades and SAT scores.

A simple regression model utilizing Principles I grades as the independent variable were run against GPAs. The results are presented in Table 5, Column A. The analysis yielded a significant result in the anticipated direction for Principles I grades with an adjusted R-Square (R2) value of .1875. The model was re-run using SAT scores as the independent variable. They too emerged as being significant (Table 5, Column B). Finally, Column C contains the results using both Principles I grades and SAT scores as independent variables. This time only Principles I is significant. Once again, even when SAT scores are introduced into the model, a student's performance in their first accounting class retains significant explanatory power.

TABLE 5
Principles of Accounting I as a Predictor of Core Grade Point Average (GPA)

Principles I	SAT Score	SAT/Principles I Both
p=.0055*	p=.0143*	Prin I p=<.0062*
R2=.1875	R2=.1433	SAT score p=.1139

***S=Signif at p<.01; **S=Signif at p<.05; ***S=Signif at p<.10; R2=R-squared (Adj)**

Finally, the researchers made an ad hoc decision that a student with a GPA of at least 3.25 out of 4.0 would be considered to have a high GPA. An indicator variable with a value of 1 was assigned to students identified as having high GPAs. A value of 0 was assigned if a student's GPA was less than 3.25. Once again a simple regression analysis specifying high GPAs as the dependent variable was conducted using Principles I grades as the independent variable. The result of this analysis is presented in Table 6, Column A. The analysis yielded a significant result

in the anticipated direction for Principles I grades with an adjusted R-Square (R2) value of .0313. The model was re-run using SAT scores as the independent variable (Column B). It too emerged as being significant. Finally, Column C contains the results for a multiple regression model using both Principles I grades and SAT scores as independent variables. This time both the Principles I grades and SAT scores emerged as significant. However, the model's R-Squared (R2) value of .0523 indicates that while statistically significant, both variables combined only explain about 5% of the variation involved in a student earning a "high GPA". Even so, a student's performance in the first accounting class retains significant explanatory power after including the effect of SAT scores.

TABLE 6
Principles of Accounting I as a Predictor of High Grade Point Average (GPA>=3.25)

Principles I	SAT Score	SAT/Principles I Both
p=<.0005* R2=.0313	p=.0018* R2=.0193	Prin I p=<.0005* SAT score p=.0410** R2=.0523

***S=Signif at p<.01; **S=Signif at p<.05; ***S=Signif at p<.10; R2=R-squared (Adj)**

DISCUSSION AND IMPLICATIONS:

The null hypothesis was rejected for seventeen of the eighteen courses which were examined. In the case of the eighteenth course, Auditing, the null hypothesis was partially rejected because, although performance in the Principles I course, as measured by final course grades, was significant in the simple regression model, it was not significant in the presence of SAT scores. The results of this research support the contention that performance in a student's first accounting course provides insights into their performance in other courses. This appears to be true whether the course is a subsequent accounting course; a non-accounting, quantitative, business course; or a non-accounting, non-quantitative business course. This result appears to be true even when a student's SAT score is taken into consideration. A student's performance in their first accounting course also appears to provide insight into their potential for successful graduation as an accounting major, as well as their final GPA. Once again, this is true even after the effects of the students' SAT scores have been considered.

The conclusion drawn from this research is that students who do well in their initial accounting course, tend to do well in their later coursework. Likewise, a student who gets off on the wrong foot, tends to have difficulty in subsequent courses, even though the courses may fall outside of the accounting discipline. A few possible explanations for these findings emerge. First, it could be that SAT scores are a proxy for a student's innate abilities but say little about his/her willingness to exert effort. It could also be that students' grades in their first accounting course provide valuable incremental information about their willingness to work hard. In essence, performance in the first accounting course is an indicator variable for a student's work ethic. This may help to explain why, many times, students who work hard, but possess lower innate intellectual abilities (SAT scores), outperform their "brighter" counterparts.

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

SUMMARY AND FUTURE RESEARCH

In all of the instances examined in this study, a significant correlation was found between performance in a student's first accounting course and performance in subsequent business courses. This result held regardless of whether the course was an accounting course; a quantitative, non-accounting business course; or a non-quantitative, non-accounting, business course. Future research needs to be conducted to determine if the course content of the introductory accounting course could be used as a construct for critical thinking abilities. It could be that the initial accounting course helps to develop important critical thinking skills which students then employ in other business courses. Future research is also necessary to determine if a student's performance in the initial accounting course could be used as a proxy for his/her work ethic. In addition, further research needs to be conducted at other universities to determine if the results of this study are institution specific, or can be generalized.

REFERENCES

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

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

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

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

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

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

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

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

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

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

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

[12] Wheeler, P., (2001). The Myers-Briggs Type Indicator and Applications to Accounting Education and Research, *Issues in Accounting Education*, 16 (1), 125-150.

INTERNAL MARKETING ESTABLISHES A CULTURE OF LEARNING ORGANIZATION

Yafang Tsai, Department of Health Policy and Management, Chung-Shan Medical University, Taiwan, (886)-4-24730022 ext.12127, avon611@gmail.com

Shih-Wang Wu, Department of Hospital and Health Care Administration, Chia Nan University of Pharmacy & Science, Taiwan, (886)-6-2664911 ext.5225, scottwu101@mail.chna.edu.tw

ABSTRACT

Internal marketing is a way to assist hospital administrators in improving the service quality provided by staff to patients. Through internal marketing, a hospital can establish a learning culture and enhance the organizational commitment of its employees. This study aims to explore the relationships between learning organization, internal marketing, and organizational commitment. This research uses a cross-sectional study to survey nursing staff perceptions regarding learning organization, internal marketing, and organizational commitment. The sample includes a medical center hospital and a regional teaching hospital in Taiwan. Two hundred questionnaires were distributed, and 114 valid questionnaires were returned, yielding a response rate of 57%. The survey process lasted from October 1 to October 31, 2009. The data were analyzed with SPSS 17.0, including descriptive statistics and inferential statistical analysis. The findings reveal that internal marketing plays a critical role in explaining employee perceptions of learning organization and organizational commitment.

Keywords: learning organization, internal marketing, organizational commitment

INTRODUCTION

In an era of knowledge-driven economies, knowledge and learning are the vigor that can not only push individuals, organizations, and societies to progress continuously but can also expand organizational development [1]. The most important assets of an organization are personnel and knowledge; a learning organization emphasizes the fact that individuals in an organization are the focal points for sharing organizational information and creating knowledge. One of the most important leadership tasks in learning organizations is to become a role model and a strong sponsor of the learning efforts of such organizations [2]. By establishing a learning organization, organizational managers can promote the sharing and creation of knowledge among different department members and promote continual learning within the organization.

Because the Bureau of National Health Insurance (NHI) in Taiwan has focused on reducing its financial deficit for several years, the predicament faced by hospitals and the health care administration is becoming more complex. As in many other countries, hospital managers in Taiwan face the challenges of controlling costs while maintaining service quality. As a result, hospital administration teams must now focus on patient satisfaction and employee organizational commitment to ensure hospital development and stability in a competitive environment [3]. Organizations can not only enhance creativity and efficiency through individual learning, but they can also satisfy employees' job demands and improve the organizational commitment of staff by promoting the idea of a learning organization. Learning organizations should create "increments of collective knowledge as a vital resource to be used internally in building agency and externally in creating competitive advantage" [4] in a competitive environment. Hospital staffs must continually

learn to provide superior service to customers; this makes hospital administrators pay more attention to the promotion of a learning organization in Taiwan.

Ballantyne [5] argues that internal marketing plays an important role in promoting learning organization, one that can improve cross-functional flows of information and tap into experience-based personal knowledge from diverse sources. Managers implement internal marketing to promote knowledge exchange among different departmental staffs and improve their skills to ensure that they can accomplish their assigned mission. Thus, internal marketing contributes to the promotion of organized learning activities. This study explored the relationship between learning organization, internal marketing and organizational commitment.

Importance of the study and research purpose

Learning organization is a management approach that can be applied to a wide range of organizations, including businesses, government agencies, and schools [6]. Health care organizations are highly knowledge-intensive institutions that require continual learning to improve their capabilities. Establishing a culture of learning organization is an important issue for health care organizations, especially for nursing care institutions [7].

Organizational commitment is a predictor of nursing turnover [8]. Employee mobility and turnover lead to higher labor costs in the form of recruitment and training. As employees become more committed to their organization, the employee retention rate improves, which in turn reduces operating costs [9]. How can one improve employees' organizational commitment? Internal marketing plays an important role by reinforcing a hospital's mission, strengthening employee relationships, increasing organizational commitment, and reducing turnover. Internal marketing is a useful tool for hospital administrators to improve employees' organizational commitment [3]. Most previous studies that have adopted an internal marketing approach have focused on for-profit sectors such as the service sector [10], airlines [11] and banks [12]. However, administrators of nonprofit organizations, such as hospitals, should pay greater attention to improving the internal management process, due to the influence of changes in the medical environment and health policy [3]. Cooper & Cronin [13] concentrated on nursing home care, and their focus was limited to internal marketing activity. However, the one goal of internal marketing is to enhance the organizational commitment of employees [14]. This study explored the relationship between learning organization, internal marketing and organizational commitment.

BACKGROUND

Learning organization

Learning organizations are able to transform themselves in the face of multiple complex forces. They are designed "deliberately to facilitate learning [4], where people are continually learning how to learn together [15]. Learning organization actions are purposive and the patterns of interaction in members' approach to work provide opportunities for learning and demonstrate responsibility and collective competence in addressing organizational goals [6]. Because learning organizations use generative learning for to conduct continuous evaluation and provide feedback to reprogram how information is processed, learning organizations encourage experimentation and system-wide thinking.

Different labels may be used to describe the processes that reflect this action component of organizational learning, such as team orientation, innovation, involvement, information flow, tolerance for error, and result orientation [16]. Senge [17] listed the five disciplines of learning organization: personal mastery, mental models, building shared vision, team learning, and systems thinking. Huw et al. [18] proposed five key features of a learning organization: open systems thinking, improving individual capabilities, team learning, updating mental models, and cohesive vision. How does one determine if an organization has the features of a learning organization?

O'Neil [19] describes four dimensions of measurement variables used to identify learning organizations: experience, structure, culture and information.

Internal marketing

The concept of internal marketing first appeared in the service marketing literature in the early 1980s [11]. Internal marketing stresses that organizations can retain employees through motivating, mobilizing, cooperating, managing, and continuing to improve the services provided to external customers [20]. It also focuses on training and inspiring employees to serve customers well [21]. Tansuhaj et al. [22] proposed that internal marketing must involve various plans that emphasize employee development and that a complete internal marketing program must necessarily include employee recruitment, training, encouragement, communication, and retention activities that lead to changes in employee attitudes and behavior. In other words, internal marketing is a company's human resource management tool used to successfully hire, train, and encourage employees and provide better services to customers [23].

Organizational commitment

The extensive body of scientific studies on organizational commitment, which largely focus on organizational behavior/psychology, has produced various definitions of the factor construct and a considerable discussion of its development, consequences, and measurement [24]. At a general level, most researchers agree that organizational commitment describes a psychological state that characterizes the relationship between an employee and the organization for which they work and that has implications for the employee's decision to remain within the organization [25]. Organizational commitment involves employees' attachment to their organization [26] [27]. Meyer & Allen [28] have identified three types of organizational commitment. Affective commitment describes individual emotional attachment and identification with and involvement in a particular organization. Furthermore, continuance commitment reflects an employee's awareness of the costs of leaving an organization. Finally, normative commitment reflects an individual's sense of obligation to remain in an organization.

Relationships among learning organization, internal marketing, and organizational commitment

The first goal that organizations attempt to achieve in a learning organization is to mold an organizational culture of learning [29]. The courses are not only meant for training or simple individual development; they aim to popularize "learning" in all aspects of the organization, thereby providing the best service [30]. Organization administrations should encourage staff dialogue [31], interaction and knowledge exchange among members of different departments and seek internal communication channels to announce policies, vision, and targets of the organization when promoting a learning organization [32]. Internal marketing encourages administrators to transmit the vision and goals of their organizations and allow the staff to clearly understand their mission by communicating with different departments. The most important mission is to change the "forced nature" of traditional training into voluntary learning. Internal marketing also emphasizes the need for the organization to select staff with passion, strengthen their capability by training [33], and nurture service-oriented behavior by clarifying their business vision and management beliefs [34]. Both of the abovementioned perspectives demand that organizations transmit the vision to their staff and adopt human resource management activities to improve their ability to serve external customers, thus creating good customer value [14]. Therefore, learning organizations have positive effects on internal marketing.

Hypothesis 1. There is a significantly positive relationship between learning organization and internal marketing.

Internal marketing is based on the idea that organizations satisfy their employees' needs through their work and are committed to their employees [35]. This commitment should be reflected by employees by demonstrating enhanced job satisfaction, loyalty, and trust [36]. Previous studies have shown that employees' organizational commitment is positively related to job satisfaction [37]. Internal marketing emphasizes the fact that organizations can hire employees who can help achieve organizational missions while providing a clear vision and goals [38]. Organizational commitment indicates the degree to which individuals are attached to an organization and identify with organizational goals [39]. Organizations can use internal marketing programs to enhance employee organizational commitment [40].

Hypothesis 2. There is a significantly positive relationship between internal marketing and organizational commitment.

A health care organization is one kind of knowledge-dense organization. Managers can ensure health care quality and technology and can then meet the requirement of patients for health care service via the staffs' continual learning. Learning organizations emphasize the fact that organization managers should empower employees with respect to their work. Health care organizations should also empower their staff by authorizing them to immediately address questions when caring for patients [42]. This can improve service efficiency for patients.

Health care staffs usually transfer knowledge, introduce new concepts, and enhance their abilities by communicating and coordinating with members of the health care team. Hospital administrators should not only encourage their staff to solve problems on the job through systematic thinking, but they should also encourage team learning. Though learning improves employees' skills, employees gain more by solving health care problems on their own, which can improve employees' commitment to an organization.

Hypothesis 3. There is a significantly positive relationship between learning organization and organizational commitment.

METHODS

Design

This research is a cross-sectional study focused on nurses of various hospitals in Taiwan.

Sampling

The respondents in this study were all nurses working in a medical center hospital and a regional teaching hospital that have implemented a learning organization culture for several years in Taiwan. Because nurses are always busy, we adopted convenience sampling to select nurses who agreed to be involved in the study. We sent out 200 questionnaires and received 114 valid questionnaires, yielding a response rate of 57%. The survey process lasted from October 1 to October 31, 2009.

The majority of the subjects were female (86.8%). The majority had been working between seven and nine years (25.4%) and had a university education (54.4%). The majority of the hospital employees were general nurses (85.1%), but most had previously worked in a medical center (61.4%). Finally, the majority was between 31 and 40 years of age (50.0%).

Data collection

Given the latent character of the variables considered in the study, multi-item Likert-type five-point scales were used (1=strongly disagree and 5=strongly agree). Learning organization (LO). We referred to the concept of a learning organization proposed by O'Neil [19] for design variables with

20 items using the Likert-type scale previously described. Internal marketing (IM). We referred to the questionnaire developed by Tsai & Tang [43] and produced 14 items using the same Likert-type scale. Organizational commitment (OC). Organizational commitment was measured by a 15-item survey according to Mowday et al. [27]; the same multi-item Likert-type scale was used.

Pilot study

First, the content of the questionnaires was developed by consulting the relevant literature. Existing items were then slightly modified to create the initial questionnaires based on the research purpose and specific industry features. Then, three directors or supervisors from the medical center were invited to perform an expert validation of the questionnaire, after which it was further revised. Next, a pilot run of the questionnaire was administered to 50 nurses. The pilot results indicated that all corrected item-total correlations were greater than 0.40 and that all Cronbach's coefficients of the subscales and scales exceeded 0.65; therefore, no item was deleted [44].

Data analysis

Descriptive statistics including the mean value (M) and standard deviation (SD) of each item and construct were reported. Exploratory factor analysis was performed to explore the factor structure, and Cronbach's α was used to evaluate the internal consistency of the factors. To examine hypotheses 1 to 3 proposed in this study, Pearson correlation analysis was used to analyze the data.

Validity and Reliability of Questionnaires

The corrected item-total correlation was used to evaluate the quality of items before performing EFA [44]. Among all three scales, two items (No. 3 and 7) of organizational commitment were deleted due to a low item-total correlation (<0.30), which indicated poor quality. Principal component analysis was used to extract major contributing factors, and Varimax rotation (Orthogonal) was performed to recognize common factors. In addition, factors with eigenvalues greater than 1 were extracted. The factor loading was greater than 0.40 was regarded as "practical significant" according to the suggestion of Hair et al. [45].

In the initial EFA of learning organization, four factors were extracted. Because the fourth dimension only dominated two items (No. 10 and 12), we decided to delete the dimension and the two items. After performing a second EFA with the remaining 18 items, three factors were extracted. The remaining three factors were named "Experience & Structure", "Culture", and "Information".
The EFA result of internal marketing indicated there were two prominent factors: one was named "Vision & Development", and the other one was named "Human Resource Management". The number of items retained was 14.

Four factors of organizational commitment were extracted. Because the fourth dimension only dominated one item (No. 15), this dimension and item were deleted. Three factors were extracted in the second EFA with the remaining 14 items and were named "Affective commitment", "Normative commitment", and "Continuance commitment". All Cronbach's coefficients of sub-construct exceeded 0.65, which is regarded as acceptable for internal consistency [44].

RESULTS

Descriptive statistics and correlation analysis

The means and standard deviations and bivariate correlations for all observational variables are given in Table 1. The mean values of learning organization ranged from 3.46 and 3.64.The mean values of internal marketing ranged between 3.48 and 3.49, whereas those of organizational commitment ranged between 3.21 and 3.46.

Table 1. Descriptive statistics and correlations for all study variables (N = 114)

No.	Variables	*M*	*SD*	1	2	3	4	5	6	7	8	9	10
1	Learning Organization(LO)	3.58	0.50	—									
2	Experience & Structure	3.64	0.52	.92***	—								
3	Culture	3.46	0.61	.88***	.68***	—							
4	Information	3.62	0.58	.82***	.62***	.69***	—						
5	Internal Marketing(IM)	3.48	0.50	.71***	.65***	.66***	.55***	—					
6	Vision & Development	3.48	0.53	.68***	.63***	.62***	.51***	.92***	—				
7	Human Resource Management	3.49	0.59	.60***	.53***	.56***	.48***	.88***	.63***	—			
8	Organization Commitment(OC)	3.37	0.41	.65***	.68***	.54***	.42***	.68***	.60***	.63***	—		
9	Affective Commitment	3.39	0.55	.55***	.56***	.45***	.40***	.58***	.52***	.52***	.78***	—	
10	Normative Commitment	3.46	0.56	.61***	.64***	.50***	.38***	.62***	.58***	.53***	.81***	.48***	—
11	Continuance Commitment	3.21	0.65	.19*	.22*	.18	.07	.21*	.12	.27**	.52***	.02	.26**

Note: * $p < .05$, * $p < .01$, *** $p < .001$.

CONCLUSION AND RELEVANCE TO CLINICAL PRACTICE

Because of extreme variations in the industrial environment, we must improve internal management skills and sustain competitive advantages. Learning organization and internal marketing are considered useful ways to succeed in this respect. For hospitals that deliver health care services to patients, the competence and skills of employees will affect the quality of service that they provide. Administrators of hospitals must improve employee competence before they can develop employees' operative skills. Setting up a learning organization is one of the ways to improve employee competence. As competencies rise, administrators should devote efforts to select and retain outstanding employees and avoid internal recruiting. Internal marketing is an important human resource management activity that can be used to maintain an excellent staff. The results indicate a positive correlation between learning organization and internal marketing and a staff's organizational commitment. We suggest that hospital administrators apply these findings to improve employee competence, increase the centripetal force between organizations and employees and enable them to face the challenges of violent variation.

This study involved only the nursing staff of two hospitals as the sample respondents. We will later increase the sample size to improve the extrapolation of results in future studies.

REFERENCES

References available upon request from Yafang Tsai.

Can the Capability of OEM Manufacturing transfer to Brand Building?

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

ABSTRACT

Taiwanese Firms have made an outstanding performance of OEM and ODM business in the PC industry. However, they are confronted with a situation in search of a next business model due to OEM and ODM no longer competitive. By investigating the development of PC industry, this paper seeks to answer this question using behavioral theory of the firm, dynamic resource-based theory and transaction cost economics, and attempts to link resources with capabilities and make contributions to the strategic management literature. First, our results would clarify the relationship between resources and capabilities building. Second, this paper enhances understanding of how firms build capacities and how to transfer it. Finally, this research examines the importance of capabilities to firm performance.

INTRODUCTION

Reviewing the history of the personal computer (PC) and Notebook industry, the OEM business has a prominent position and successful outsourcing strategy provides an exemplar of an international division of labor in value chain activity. Major American computer companies have considerably expanded their OEM purchases from Taiwanese sources. IBM sources monitors from Philips, Taiwan and Sampo; motherboards from GVC, Elite and Lung Hwa; power suppliers from Delta and Sun-Moon-Star and laptops from ASE (Ernst, 2000). Indeed, Taiwan has become the world's largest producer of notebook computers, as well as a range of related-PC components including keyboards, motherboards, monitors, scanners, and power supplies

Over the past few decades, Taiwan's manufacturing industry had focused on OEM and ODM outsourcing businesses. The orders from international buyers were mainly for OEM/ODM assembly and production, and most of these buyers based their purchase decisions on product specifications, functions and prices. Thus, there is little room for Taiwanese design companies to add value to products on branding.

With the rise of China and Southeast Asia countries, however, Taiwan no longer has competitive edges in OEM/ODM production. Many Taiwan-based enterprises have begun to establish their own brand manufacturing, OBM, and as a result, design has become a powerful tool to enhance product features. Through the design of stylish appearance, structure, function and operation, higher product value would be created

Given these trends, here are some questions which researchers and practitioners should ask: Why not use own brand name in the market if you can produce such a good quality product? Why can't most of the OEM suppliers survive? Why do only a small number of OEM suppliers exist? Can OEM suppliers change their strategies into own brand? How to transform existing resources into capabilities and obtain sustainable competitive advantage? Amit and Schoemaker (1993) maintain while empirical models may, ex post, point to a limited set of resources and capabilities that explain some of the firm's past performance, ex ante such models offer limited insight into the dimensions of competition that will prevail in the future. For managers, the challenge is to identify, develop, protect, and deploy resources and capabilities in a way that provides the firm with a sustainable competitive advantage and, thereby, a superior return on capital.

Efficiency (manufacturing)

Taiwan is a place with no names, at least no brand names that ring familiar in Western ears. Japan,

of course, has plenty of names that thunder across the world, from Hitachi and Honda to Toshiba and Toyota. But Taiwan has long been an island where workers anonymously toil to produce components and often whole computers that other people put their famous company names on (Kraar, 1995). Most of the Taiwanese computer companies devote to making a few components on which their very survival depends. They have elevated to a fine art such seemingly workaday tasks as controlling inventory and overhead expenses. Everyone competes for management efficiency. Dell Computer in the U.S. considers itself lean with overheads equivalent to about 15% of sales, but by comparison, overheads of Taiwan computer companies appear almost anorexic, averaging less than 10% of sales. A fascinating Taiwan money-saving technique: 'modular manufacturing,' in which price-sensitive parts are left out of PCs and then shipped to distributors for installation at the last possible moment before the machines are sold (Kraar, 1994)

Many Taiwanese companies are OEM/ODM suppliers, designing and manufacturing machines, then selling them to the big-brand name players. The Taiwanese OEM/ODM suppliers have tiny margins. Their customers, Dell and HP of the world, are not about to allow them to widen those margins. If your are a Taiwanese tech executive, you have no pricing power (Einhorn, 2003)

Taiwanese companies are making serious headway, racking up design prizes and gaining market share over foreign competitors. Some companies, such as BenQ, aim to emulate the success of South Korea's Samsung Electronics, using sharp design to leverage their contract manufacturing business into a global brand. Acer, BenQ's former parent, has surpassed big names like IBM and Toshiba to become the second-best-selling notebook PC brand in Europe, after HP. Others, such as Quanta Company and Lite-On Technology, while not pushing brands, are pouring resources onto product innovation to win more profitable contracts from their customers.

Taiwan companies still have plenty of distance to go to catch design leaders like Samsung, Sony and Apple Computer. Another problem for Taiwan companies to build their own brand is to take risk of competing with their contract customers. BenQ was the biggest outside supplier of mobile phones to Motorola, but the U.S. company has shifted orders elsewhere as BenQ's brand profile has risen. Asustek, wary of alienating its contract customers, does relatively little advertising for its PCs (Dean, 2002).

With the exception of Acer, most major Taiwanese companies do not have their own brand-name businesses. They are equipment providers, not selling directly to consumers but to companies such as Dell and IBM (Einhorn, 2002).

Due to high cost and high failure rate in brand building, manufacturing is preferable in terms of short term efficiency. The recent example is BenQ, once one of Taiwan's top consumer-electronics brands. In 2005 management bought the money-losing cellular division of giant Siemens and paired the BenQ and Siemens name on phones. The German subsidiary was liquidated, and now BenQ, renamed Qisda, focuses on contract manufacturing (Einhorn, 2007).

Due to the characteristics of OEM business, most of firms have to 'lock in' specific investment and organizational routine and make endeavors to pursue efficiency for their large quantity purchase order. Since PC is already in the stage of commodity, price and delivery become the major competition factors. OEM firms will loose purchase order if they are not best in efficiency.

Firm Performance

The commodification of manufacturing as an outcome of technological advances in codifying the interface between functions in the production system poses a particular challenge for Taiwanese firms. Manufacturing capabilities have been the core asset of Taiwan. Taiwan's strengths are in manufacturing, not services, in quick technology absorption, competition on price, R&D for process improvements, and low-cost inputs. As the Taiwanese face more and tougher competition today on manufacturing, they have two broad options going forward. On one hand, strategies that center on

competing on price and lowering costs, and, on the other hand, strategies that use innovation, differentiation, or branding in order to build barriers to entry and to escape competitive pressures (Berger & Lester, 2005).

The most serious problem of relying on a strategy of lowering costs is that this is a game that many other can, and do, play. And as they do, the margins for OEM manufacturers shrink. The competitive pressures are intense in China as manufacturers concentrate on driving down costs across a range of products with very low barriers to entry. In a world in which everyone is chasing after the same markets – air conditioners or television sets or mobile phones – price plummet, margins collapse, and soon no one is making profits and the search for yet another product in relatively easy reach begins all over again. The Taiwanese are heavily involved in this scramble to reduce costs and prices through moving operations to China, but the ultimate consequence may be very costly (Berger & Lester, 2005).

Taiwan's OEM and ODM companies have always faced the challenge of defending their profits in relationships with lead firms that are larger and more powerful. This effect is exacerbated by high dependence on a small number of customers such as HP/Compaq, Dell, Sony, Apple, and IBM, which accounted for 63 percent of all IT hardware made by Taiwanese firms in 2002. In order to reduce dependence and to escape narrowing margins and cutthroat competition, businesses need to offer unique or difficult-to-replace products and services that raise the entry bars for would-be rivals. Brand names are one way to establish such barriers. But thus far Taiwanese firms have not been greatly successful in branding. Many of the brands that have been developed by Taiwanese firms are mainly sold in Taiwan – a market that is too small to offer solid growth. But fundamental issues remain: first, can an OEM or ODM firm develop branding without doing great, even fatal, damage to relationships with its old lead firm customers? Acer's difficult experiences in trying to build a brand for its PCs while main- taining a parallel OEM business making PCs for other brands demonstrated how fierce the resistance o the lead customers can be to competition from their suppliers and how real the dangers of destroying the relationship. The lesson many of the OEM and ODM firms have drawn is that branding is too risky for their main business (Berger & Lester, 2005).

Aside from branding and innovation, there are a variety of other strategies that the companies are implementing. For companies that are good at incremental innovation and fast followership, diversification often seems a promising move: at a best, it is a way of offering one-stop shopping to customers. But the capabilities that Taiwan OEM and ODM firms have developed in detailed design and production quality and speed are not easily adaptable to wide range of electronics products beyond personal computers. Diversification thus often seems simply to move firms onto nearby product niches that are soon themselves under heavy competitive pressures (Berger & Lester, 2005).

Flexibility (Adaptability)

The organization theorist Chester Barnard maintained that adaptation was the central problem of organization. Barnard was concerned with cooperative adaptation of an intentional kind. Formal organization, especially hierarchy, was the instrument through which the conscious, deliberate, purposeful cooperation to which Barnard called attention was accomplished. Transaction cost economics (i) concurs that adaptation is te central problem of economic organization; (ii) regards adaptations of both autonomous and cooperative kinds as important; (iii) maintains that whether adaptations to disturbances ought to be predominantly autonomous, cooperative, or a mixture thereof varies with the attributes of the transactions (especially on the degree to which the investments associated with successive stages of activity are bilaterally or multilaterally dependent); and (iv) maintains that each generic form of governance – market, hybrid, and hierarchy – differ systematically in its capacity to adapt in autonomous and cooperative ways. A series of predicted (transaction cost economizing) alignments between transactions and governance structures thereby obtain, which predictions invite and have been subjected to empirical testing (Williamson, 1993).

When a firm's external or internal environment changes, routines (which include decision rules) need not necessarily remain constant. 'Problem-solving responses routinely evoked by difficulties with existing routines' may lead to a change in routines. As an example in Nelson and Winter (1982), changes in prices may cause firms to search for an alternative to the current decision rule for responding to price changes. This example suggests that history-dependent firm behavior may exhibit varying degrees of adaptation or rigidity. The extent of change in an activity depends on: the amount and type of change in the firm's external or internal environment; the specific decision rule that guides a firm's alterations in activity in response to environmental change; the firm's specific routinized problem-solving responses to difficulty with current routines (search effects) that may produce changes in routines or decision rules; and ultimately, feedback from the environment once the firm makes initial changes in the activity (selection effects) (Helfat, 1998).

In the face of a lead firm's desire to maintain both flexibility and control, a supplier is often forced to upgrade its manufacturing competence to ensure its competitiveness. There are risks associated with continuous investments in manufacturing, especially if they entail relation-specific investments and lead to long-term captive supply relationships. But modular products and processes provide suppliers with opportunities to reduce asset specificity while maintaining their capability to offer flexible manufacturing services. Having put such generic capabilities in place, a supplier can further expand its customer base and serve new application areas, as long as the processes are applicable to those areas. In other words, given the right mix of capabilities, a supplier can reduce its risk through structuring a multiple-customer and/or multiple-business profile (Sturgeon & Lee, 2005).

Firm may search for other opportunity in addition to OEM business. However, the switching cost is high and it is not easy for OEM supplier to take turnaround strategy. In the process of re-positioning, adaptability may be a feasible way to fulfill their brand strategy.

Acer spun off its design, manufacturing and services unit into a newly independent company called Wistron. The move signaled Acer's desire to make good on a promise to shake up and streamline its organization structure. The main purpose of the separation is to avoid the appearance of potential conflict between OEM and Acer-brand business efforts. Stan Shih said: "In the future, the new Acer will become a marketing and service company, while Wistron will remain focused on OEM.'

By spinning off Wistron, Acer's brand-name operations will also be freed up to seek more competitive contract manufacturers, thus driving down costs. What's more, Wistron will no longer be under an obligation to produce Acer-brand products. Perhaps the most significant benefit will be the end of conflict-of-interest issue, in which the company was sometimes seen as competing against its OEM buyers, such as IBM, with its own brand of PC products (Clendenin, 2001).

The concentration of ODM firms in PC industry, and its dependence on several extremely large buyers such as Dell and HP/Compaq, comes with several potential threats and challenges placed a great deal o pressure on ODMs to lower costs. The second risk comes from the fact that spending on PCs accounts for a shrinking portion of overall IT spending worldwide. These risks highlight an urgent need for Taiwanese ODMs to diversify their product and customer scope beyond PC-related products, and major ODM firms have sought to do so, with some success. ASUSTek, a leader in motherboard design and production, has successfully diversified into notebook computers, VGA cards, optical disk drives, and the manufacturing of the PS2 video game console for Sony. Revenues have increased dramatically, while motherboard sales have dropped to only one-third of total sales. ODM firms such as Hon Hai (Foxconn), Quanta, Compal, and ASUSTeK have allocated significant resources to the design and production of mobile-phone handsets. Mobile-phone handsets provide a logical migration path for ODMs because, like personal computers, they have rapid postarchitectural design cycles based in standard chipsets, and like notebook computers require industrial design expertise related to miniaturization. However, because neither detailed design competencies nor the PC-specific supply base in Taiwan are easily adaptable to a wide range of electronics products beyond PCs, product and customer diversification has been slow and the ODM

firms remain quite narrowly focused on PCs, mobile phones, and similar products. Still, except for the near monopolies held by Intel and Microsoft in CPUs and operating software, severe downward pressure on prices and margins is felt at all levels of the PC value chain: lead firms, contracts manufacturers, and components suppliers alike (Sturgeon & Lee, 2005).

Research Framework and Proposition Development

The research framework is shown in Figure 3. The dimensions of resources are classified into goal incongruence, uncertainty avoidance, problemistic search, asset specificity, asymmetric information, knowledge and reputation spillover, organizational routine, isolating mechanism and high growth rate in related businesses. These dimensions are expected to affect the choice of capabilities building. Both exploration and exploitation of capabilities building are positively related to firm performance. Organizational researchers have maintained for multiple dimensions of organizational performance such as effectiveness and efficiency (Auh & Menguc, 2005). In this regard, we propose that firm performance can adopt an effectiveness perspective (e.g., brand management), efficiency perspective (e.g., manufacturing) or flexibility (e.g., adaptation).

There are often differences stemming from the political nature of an organization. Goals are ambiguous, and commitment to them is confounded by their relation to personal and subgroup objectives. Conflict and decision advocacy within putatively rational decision processes lead to inflated expectations and problems of implementation and thus to disappointments. Different groups in and organization often have different targets and evaluate the same outcome differently. New organizational leaders are inclined to define previous outcomes more negatively than ore the leaders who preceded them. As a result, evaluations of outcomes are likely to be more negative or more mixed in organizations than they are in individuals.

Organizational success is ordinarily defined in terms of the relation between performance outcomes and targets. Targets, however, change over time in two ways. First, the indicators of success are modified. Accounting definitions change; social and policy indicators are redefined. Second, levels of aspiration with respect to any particular indicator change. The most common assumption is that a target is a function of some kind of moving average of past achievement, the gap between past achievement and past targets, or the rate of change of either (Levitt & March, 1988).

In keeping with virtually all theories of organizations, we assume that the coalition represented in an organization is a coalition of members having different goals. We require some procedure for resolving such conflict. Basically, we have maintained that most organizations most of the time exist and thrive with considerable latent conflict of goals. Except at the level of non-operational objectives, there is no internal consensus (Nelson & Winter, 1982).

Cyert and March (1963) maintain that individual participants in the organization may have substantial different preference ordering and actual organizational goals cannot normally be described in terms of a joint preference. The relation between separate goals may even result in conflict. Take production and sales goal for instance, the production managers usually prefer orders with large quantity while managers of sales division strive for orders even only with small quantity. Large quantity is easy in planning for material purchase, production scheduling and efficiency improvement. Actually, large orders also originate from small orders. The small order including prototype should also list in a higher priority. However, large orders always become the first priority due to cost reduction, yield rate and short-term efficiency considerations. We therefore propose:

P1: As the level of goal congruence increases, a firm is more likely to choose exploitation rather than exploration for capability building

Reasoning: Due to the goal congruence between short-term vs. long-term, capacity utilization vs. profit, operational managers vs. executives, firms are more likely to choose exploitation such as solicit purchase orders to fully utilize the factory capacity which might

sacrifice profit.

In the case of the business firm, there are uncertainties with respect to the behavior of the market, the deliveries of suppliers, the attitudes of shareholders, the behavior of competitors, the future actions of governmental agencies, and so on. As a result, much of modern decision theory has been concerned with the problems of decision making under risk and uncertainty. The solutions involved have been largely procedures for finding certainty equivalents or introducing rules for living with the uncertainties.

Our studies indicate quite a different strategy on the part of organizations. Organizations avoid uncertainty: (1) They avoid the requirement that they correctly anticipate events in the distant future by using decision rules emphasizing short-run reaction to short-run feedback rather than anticipation of ling-run uncertain events. They solve pressing problems rather than develop long-run strategies. (2) They avoid the requirement that they anticipate future reactions of other parts of their environment by arranging a negotiated environment. They impose plans, standard operating procedures, industry tradition, and uncertainty-absorbing contracts on that environment (Nelson & Winter, 1982).

Pindyck (1993) discusses in detail the implications for investment decisions of two types of cost uncertainty for projects (e.g., a nuclear power plant) that take time to complete. The first type of cost uncertainty is technical uncertainty, i.e., uncertainty over the physical difficulty of completing a project, such as how much time, effort, and materials will ultimately be required for completing the project. Such uncertainty is only resolved as the investment proceeds but is largely diversifiable. The second type of cost uncertainty is input cost uncertainty, i.e., uncertainty over the prices of construction inputs or over government regulations affecting construction costs. Such uncertainty is external to the firm and may be partly non-diversifiable. With regard to reduction of endogenous uncertainty through organizational actins, McGrath (1997) suggests that each firm has its own uncertainty profile in technology development and commercialization, and that firms can make pre-amplifying investments to influence uncertainty to their advantage and enhance the economic value or the ‘appropriability’ of the value of technology options (Li, James Madhavan & Mahoney, 2007).

In order to reduce the technology uncertainty, OEM suppliers keep a good relationship with OEM buyer through joint research in product design of white box, gray box or black box which indicate the percentage of design drawings OEM buyers will provide. As for rapid fluctuation of price in components and materials, OEM buyers may request to prepare key components such as CPU for their suppliers through central purchasing in order to ensure delivery and cost reduction. We therefore propose:

P2: As the level of uncertainty avoidance increases, a firm is more likely to choose exploitation than exploration for capability building

Reasoning: OEM buyers provide one year forecast and place orders in month or quarter base. New models for future order are offered by RFI (request for inquiry) and RFQ (request for quotation) continuously. Existing OEM orders in production usually maintain in one or two years.

P3: As the level of organizational routine increases, a firm is more likely to choose exploitation over exploration for capability building

Reasoning: When purchase order is set, the production of one model may continue more than one year. The process from RFI (request for inquiry), RFQ (request for quotation), order processing, production, QC (quality control), delivery become routine and firms are reluctant to change.

A STUDY OF THE IMPACTS OF SOCIAL CAPITAL ON LEARNING ACTIVITIES WITHIN COMMUNITIES OF PRACTICE (COPS)

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

Ping-Chuan Lee, Dept. of Business Administration, Chung Yuan Christian University, Taiwan, R.O.C.
E-mail: supercute.apple@gmail.com

ABSTRACT

This study focuses on the relationship between social capital and learning activities of the CoPs along with the effects of learning activities of the CoPs on the R&D performance. We adopt a three-dimensional social capital consisting of interaction (structural dimension), trust (relational dimension) and perspective taking (cognitive dimension); and a four-dimensional variable of learning activities of the CoPs that includes observation, conversation, practicing and bargaining. Using a questionnaire designed to test the research hypotheses, the study surveyed 300 participants from R&D project teams of firms belonging to the high-tech industries in Taiwan, and received 204 effective questionnaires (a respondent rate of 68%). The results find that: (1) Among the three constructs of social capital, trust and perspective taking have positive impacts on all four constructs of learning activities of CoPs, while interaction largely has no effects on learning activities of CoPs except for a mild effect on bargaining. (2) Among the four constructs of learning activities of the CoPs, both observation and bargaining positively affect R&D performance, while conversation and practicing have no significant impact on R&D performance.

Keywords: Communities-of-practice (CoPs), Social capital, R&D performance

INTRODUCTION

In the modern economy, the intangible knowledge may be more important than tangible assets such as natural resources in maintaining a firm's competitive advantage. Because intangible knowledge is difficult to imitate, so firms can bring sustainable competitive advantage. Peter Drucker (1995) even considers that knowledge is the only meaningful resource. To enhance a firm's performance requires carefully managing the firm's knowledge base aggregating from its all members' knowledge pool. Thus, the personal knowledge of an individual that is accumulated through a lifetime of experience, experimentation, perceptual observation, and learning forms an important foundation for a firm performance. To enhance a firm's level of innovation and research and development (R&D) outcomes requires an organization's members continuously enlarge their learning activities (Mascitelli, 2000; Pitt and Clarke, 1999).

Learning is an inherent part of an individual's participation in work practice in the community. Because learning is a matter of engagement in socially defined practices, the communities that share these practices play an important role in shaping learning (Wenger, 1996). Communities of practice (CoPs) as an emergent form of organizing become more and more important of the knowledge management practice in today's knowledge economy. Forming through the networks connecting people who engage in implementing work practice of the same knowledge domain, CoPs promote practitioners' interaction with each other and significantly stimulate the sharing of knowledge between practitioners and will give the organization a competitive advantage and efficiency. Wenger,

McDermott, and Snyder (2002, p.12) argue that organizations need to cultivate communities of practice actively and systematically, for the benefit of organizations as well as for the benefit of the members of the communities themselves.

Social capital refers to the resources individuals within a community can draw upon to provide value to themselves and their organizations. By stressing the social basis of cognition in which people organize their thinking and action about innovation, social capital is fostered and tapped whenever needed to enable individuals to perform their job more efficiently and effectively, even enable an organization to better manage its knowledge resources (Lesser and Prusak, 1999). Communities that are rich in social capital tend to communicate well. Their members spend time together, help each other, and contribute to the collective common good. And, as mentioned by Lesser and Storck (2001), the social capital resident in CoPs leads to behavioral changes which in turn may result in business advantage. Thus, CoPs can be the "engines for development of social capital" (E. Invernizzi, S. Mauriello and S. Romenti, 2010). Hence, a main purpose of this study is focused on this social capital's effects on the learning activities of CoPs.

Although R&D performance is always the greatest concern of most networks of teams in R&D projects define the opportunities potentially available to create new knowledge (Uzzi, 1997; Sosa et al., 2004), learning activities of CoPs is used as a mediating variable between the social capital and R&D performance because learning activities is more suitable than R&D performance to receive the direct consequence exerting by the CoPs' social capital and can more directly exhibit the synergistic effect of teamwork.

Research framework

To sum up, this study takes CoPs into research model, and tries to verify the relationships between social capital, learning activities of the CoPs and R&D performance as depicted in Figure 1. Accordingly, the purposes of this research are: (1) To explore that the relationship between social capital and learning activities of the CoPs. (2) To explore the relationship between learning activities of the CoPs and R&D performance.

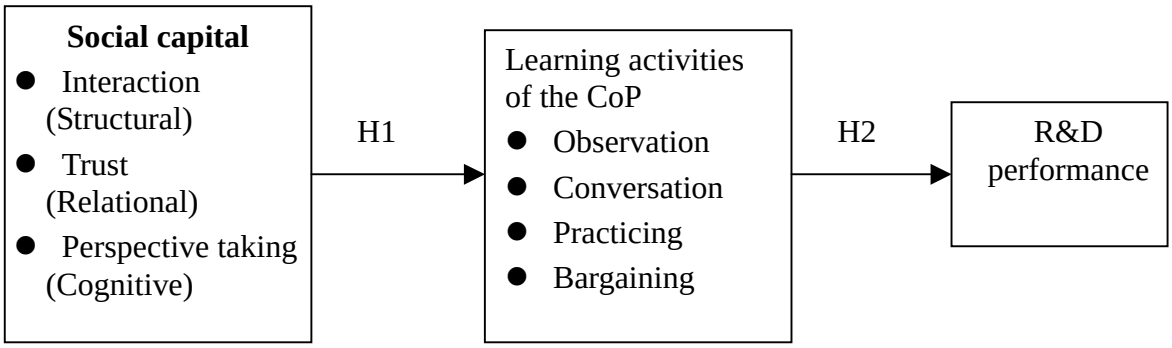


Figure 1 Conceptual Framework

LITERATURE REVIEWS

Communities of practice

Communities of practice are defined by Wenger et al (2002) as “a group of people who come together to learn from each other by sharing knowledge and experiences about the activities in which they are engaged. Individuals, the community and organizations can all benefit from this process.”
It’s a group of professionals who interact with each other within an organization, across organizational units, or even organizational boundaries have a common interest or field of application in certain work-related topics and share their knowledge on a regular basis (Andriessen, 2005; Lave and Wenger, 1991; Wenger, 1998; Wenger, McDermott, and Snyder, 2002). A community of practice is a knowledge sharing forum for practitioners of a discipline or topic, or those interested in addressing a specific concern. In pursuit of common goals and interests, these groups employ common practices, tools, language, beliefs and values. As people work together, they learn from doing and develop a shared sense of what has to happen to get things done. They develop a common way of thinking and talking and eventually share a sense of mutual identity.

Learning activities of the CoPs

CoPs are learning environments defined as “groups of people who share a common set of problems, or a passion about a topic, and deepen their knowledge and expertise in this area by interacting on an ongoing basis” (Wenger, McDermott, and Snyder, 2002). Wilson, Goodman and Cronin (2007) argued that individual can learn within the context of a group. If learning is a matter of engagement in socially defined practices, the communities that share these practices play an important role in shaping learning (Wenger, 1996). The CoP framework involves rethinking the focus of learning from an individual mind to a process that unfolds in a participatory framework (Lave and Wenger, 1991). Learning is viewed as distributed among the participants of the community, including experts, novices, and everyone in-between. Understanding and experience are in constant interaction. And CoPs usually took groups of professionals as their research subjects such as technicians (Orr, 1990), tailors and midwives (Lave and Wenger, 1991). In this study it can be seen that the theory of CoPs focuses on learning within professional groups. The important core theme of CoPs is how a person who called novice learns to become an expert through interaction within an emphasis on the loci of that learning process (Lave and Wenger, 1991).

In this study, there are four learning activities of the CoPs including observation, conversation, practicing and bargaining, as detailed in the following.

- **Observation**

CoPs are considered to be a type of learning community (Lave and Wenger, 1991). Social learning theorists suggest that communities provide a foundation for sharing knowledge. It is believed that individuals can learn by observing and modeling other people. Bandura (1977) emphasizes that observing other people's behavior allows for a safer and more efficient way of acquiring complex behaviors or skills than learning by trial and error. It provides a safe environment for individuals to engage in learning through observation and interaction with experts and through discussion. Just like the works allocated to low-status members are more peripheral than those to high-status members, but these low-status members usually can learn something through observing work practice of high-status members. Thus, on work practice, observation is also an important kind of participation in CoPs.

- **Conversation**

CoPs can provide for sharing knowledge that requires interaction and informal learning processes such as storytelling, conversation, coaching, and apprenticeship. And it also can let the people

participate in group discussions, one on one conversations, learn new ideas, or see specialists at the first battle on the side of the issue, and even bring and to promote vitality. In the potential of conversations between inside and outside, through the perspective of an outsider to help members see the opportunities, and would conversation with the community leaders and core members, the results of this dialogue, it would let CoPs which understand this issue and the legitimacy of the people, who have new opportunities, and efficient to do the seeds of change. Communities are, in many cases, about "knowing a man who can", and they enable conversations that transfer knowledge around the group (Pierce, 2002). Lave and Wenger (1991) describe that CoPs for troubleshooting would develop a shared body of knowledge based on stories and anecdote that circulated within the communities, this situation is much like apprenticeship learning which is also "supported by conversations and stories about problematic and especially difficult cases (p. 108)". Thus, a community of practice (CoP) is a set of people informally bound together through common interest and language with the goals of open conversation, and exchange and retention of pertinent knowledge.

■ Practicing

Learning is the process of becoming a member of the CoPs. The motivation to learn is the motivation to become a member (Galagan, 1993). In the CoPs the newcomer is not only a learner, learning in situation, but also an active participant in the social community. As a member of a social structure the newcomer is accepted and "legitimate", even though being located in the peripheral zone. Lave and Wenger (1991) term it "situated legitimate peripheral participation," because learning must occur in the context where it will be used (situated), allowed to occur (legitimate), and new members learn from watching and practicing from the edges of the group (peripheral participation). This situated legitimate peripheral participation occurs through several mechanisms. Stamps (1997) cited was that the way people actually acquire skills is by practicing, not by being lectured at. This idea is at the core of the CoPs viewpoint as learning occurs through legitimate peripheral practice. For example, Lave (1992) illustrates the CoPs view via Yucatec Mayan midwifery. Young Mayan girls watch as their elder relatives assist births. Eventually, she is allowed to help with the easier tasks, gradually watching, learning, and practicing increasingly difficult techniques. Ultimately, she becomes an experienced midwife, and assists the training of other girls. Thus, in the CoPs there are one activities of learning which it would through practicing to acquire the skills and expertise.

■ Bargaining

A CoP can be defined as an informal group that share an interest and from their own will, to resolve common problems or issues through bargaining their activities (Vanasse and Poynton, 2003). CoPs grow out of a convergent interplay of competence and experience that involves mutual engagement. They offer an opportunity to bargaining competence through an experience of direct participation. If the members of an organization do not actively participate, the community will be difficult to deploy its members' resources. The developments of CoPs are focus on luring and cultivating the participation, rather than planning, guidance, or arranging their activities. This process is the need to bargaining. So the members of CoPs can be more enthusiasm and sense of identity to participate the activities (Wenger, McDermott, and Snyder, 2002). Within the communities, news travels fast; community knowledge is readily available to community members. But the communities must function within corporations that treat information as a commodity and that have superior bargaining power in negotiating the terms of exchange (Brown and Duguid, 1991).

The problematic parts of learning in CoPs are rooted on their hierarchical structure. Conversations are smoother among members with similar statuses than those among members with status distance. Bargains among members are surely more difficult than conversations. For these reasons, the change

of knowledge or practice in CoPs is quite slow. Thus, according to the above description we would discern that learning activities of the CoPs have observation, conversation, practicing, and bargaining.

Social Capital

The term "social capital" initially appeared in community studies, highlighting the central importance— for the survival and functioning of city neighborhoods—of the networks of strong, crosscutting personal relationships developed over time that provide the basis for trust, cooperation, and collective action in such communities (Jacobs, 1965). Nahapiet and Ghoshal (1998) defined social capital as the sum of actual and potential resources embedded within, available through and derived from the network of relationships possessed by an individual or social unit. A central proposition of social capital theory is that networks of relationships constitute a valuable resource for the conduct of social affairs and provide their members with the collectively owned capital (Adler and Kwon, 2002). Narayan and Pritchett (1997) suggested that communities with high social capital have frequent interaction, which in turn cultivates norms of reciprocity through which learners become more willing to help one another, and which improve coordination and dissemination of information and knowledge sharing.

Nahapiet and Ghoshal (1998) used a three-dimensional model, which includes structural dimension, cognitive dimension, and relational dimension, to elucidate “the ways in which social capital influences the development of intellectual capital”. This study will borrow this social capital model to derive a framework to incorporate interaction, trust, and perspective taking as the most important factors in the structural, relational, and cognitive dimensions of social capital, respectively. The rationale of this framework is detailed in the following section.

- **The structural dimension**

The structural dimension is regarded as a feature of its network configuration. It depicts how individuals interact with others on task allocation, responsibilities, and authority (Harris and Lambert, 1998; Stewart and Barrick, 2000) as well as describes the network ties between them (Scott, 1991). It focuses on the connections and links that members have with one another. In this regard, CoPs are fine venues that provide the opportunities for individuals to develop a network of individuals who have similar interests. This manifests itself in several ways. First, the community serves as an intra-network clearinghouse by identifying those with relevant knowledge and helping individuals within the community make connections with one another. Second, the community acts as a reference mechanism, quickly enabling individuals to evaluate the knowledge of other members without having to contact each individual within the network. Lastly, the community of practice can help connect individuals from outside the network to those who are already identified as community members (Lesser and Prusak, 1999). The perceived gain for members of the community is that it will take less time and effort to gather information via the network of connections than on their own. Increased connections increase the likelihood of benefit for the community (Vanasse and Poynton, 2009).

- **The relational dimension**

The relational dimension is related as characteristics of the relationship among the organization’s members, such as trust, norms, obligations, and identification (Nahapiet and Ghoshal, 1998). Among these characteristics, trust is the most important one as Nahapiet and Ghoshal’s (1998) stated that “collective trust may become a potent form of “expectational asset” that group members can rely on more generally to help solve problems of cooperation and coordination”. In this study, trust is used to

represent the level of relational dimension of the organization's social capital. It refers to the bonds and the sense of trust that the members of the community develop for each other. Without these, members risk withholding information instead of sharing it (Vanasse and Poynton, 2009). CoPs foster the interpersonal interactions necessary to build a sense of trust and obligations critical to building social capital. By being able to bring people together to create and share relevant knowledge, the community creates the condition where individuals can "test" the trustworthiness and commitment of other community members (Lesser and Prusak, 1999).

- **The cognitive dimension**

The cognitive dimension of social capital refers to resources that are available to an individual from having shared expectations as team members interpret cues in a similar manner, make compatible decisions, and coordinate their actions (Levesque, Wilson, and Wholey, 2001; Nahapiet and Ghoshal, 1998). The cognitive dimension is an important aspect of social capital, and according to
Boisot (1995) and Boland and Tenaski (1995) argued that the cognitive dimension is predicated on meaningful communication as an essential part of combination processes. They argue that this requires at least some sharing of context between the parties involves in the exchange. Sharing might come about through shared codes and language (i.e., the extent to which people share a common language facilitates their ability to gain access to others and information). Differences in language and codes keep people apart (Nahapiet and Ghoshal, 1998), and shared narratives adhere people (Orr, 1990; Wenger 1998). Therefore, in this study, perspective taking is deemed as the main facet of cognitive dimension of social capital. It is related to the common interest or the shared understanding of the issues the community wishes to resolve.

R&D Performance

According to Brown & Svenson's (1998) R&D performance measurement system, there are two types of R&D performance. One is e efficiency of R&D process such as the implementation of listed tasks, accordance with time schedule, and/or aligned budget; the other is focused on the effectiveness of R&D results such as customers' satisfaction with the new products or services, accordance with pre-set goals, numbers of new patents. There are many literatures about the measurement of R&D performance, followings are some descriptions of them.

Kerssens-van Drongnelen and Cook (1997) used Balanced Scorecard developed by Kaplan and Norton (1996) to measure R&D performance that developed the set of measurement metrics through financial perspective, customer perspective, innovation and learning perspective and internal business perspective for measuring overall R&D performance. On the basis of the results asserted by Griffin and Page (1993), Griffin and Page provide three project-level measures for new product development success and failure through the set of recommended measures including, seven customer-based, four financial and seven technical performance measures. Loch et al. (1996) developed a two-step model to measure the performance of the NPD function confined to the electronics industry in their study. In the model, they divided development performance into two dimensions — development output performance and development process performance. Development output performance driven by development process performance — the operational management of development projects—is the direct driver of business success.

HYPOTHESES

The relationship between social capital and learning activities of the CoPs

A community of practice (CoP) is for learning from interactions with people who share the same interest. The focus of group interactions is to share and listen to other community members. A strong community fosters interactions and relationships and it encourages a willingness to share ideas, expose one's ignorance, ask difficult questions, and listen carefully. Farris (1969) and Allen (1977) found that interactions with others and increased communication have positive impacts on scientific output. Mascitelli's (2000) study indicates that more interaction among participants of an R&D project has a cumulative effect on prompting knowledge sharing, especially for tacit knowledge. Jassawalla and Sashittal (1998) also find a positive relation between interaction and collaboration that learners with high level of interactions are more willing to help others and improve coordination and dissemination of information and knowledge sharing.

A strong community fosters interactions and relationships based on mutual respect and trust. Where trust and friendship levels are high, people are more willing to engage in social exchange and cooperative interactions, such as confiding with another, asking for help, having spontaneous conversations and unstructured meetings, as well as sharing information, knowledge, and resources (Bouty, 2000; Krackhardt, 1992; Nahapiet and Ghoshal, 1998). And, as mentioned by Williams (2001), "interpersonal trust is an important social resource that can facilitate cooperation and enable coordinated social interactions". Hence, high levels of trust in relationships among members of a CoP will help create a knowledge sharing environment that results in high levels of members' learning activities such as observing, listening, practicing, questioning, debating, and collaborating.

Thus, perspective taking can help people "adequately imagine what others will do" (Dougherty, 2001) and facilitate the formation of "some sharing of context between the parties" (Nahapiet and Ghoshal, 1998). Moreover, perspective taking may result in both the cognitive effect of mutual understanding and the behavioral effect of empathic concerns towards others (Parker and Axtell, 2001; Settoon and Mossholder, 2002). Hence, perspective taking will help in developing and maintaining an agreed-upon set of terminology, codes, and narratives used by communities' members to share knowledge.

Thus, the following hypothesis is proposed:

H1: Social capital will affect learning activities of the CoPs.

- ***H1-1****: social capital will positively influence observation aspect of learning activities of the CoPs.*
- ***H1-2****: social capital will positively influence conversation aspect of learning activities of the CoPs.*
- ***H1-3****: social capital will positively influence practicing aspect of learning activities of the CoPs.*
- ***H1-4****: social capital will positively influence bargaining aspect of learning activities of the CoPs.*

The relationship between learning activities of the CoPs and R&D performance

In R&D teams, high degrees of R&D performance strongly depend on the timely delivery of new knowledge. Learning ability and knowledge application ability will reinforce team members' knowledge creation ability to obtain the outcomes of R&D (Cohen and Liventhal, 1990). It will further set a more high level of R&D goals, and actively pay attention to environmental opportunities that may arise in order to keep up with the progress of the environment, then induced the efforts of the next wave of R&D and in the courses improve the R&D performance again. Meanwhile, knowledge creation capacity of members of the CoPs is an important capability for organization to develop core competencies and products, thus enhancing R&D performance (Coombs, 1996; Hill and Jones, 1998). So when the members of R&D teams exhibit high level of learning acivities, the

R&D performance will be enhanced.

H2: Learning activities of the CoPs are positively related to R&D performance.

H2-1*: Observation is positively related to R&D performance.*

H2-2*: Conversation is positively related to R&D performance.*

H2-3*: Practicing is positively related to R&D performance.*

H2-4*: Bargaining is positively related to R&D performance.*

THE SURVEY

A questionnaire survey was designed to explore the relationship among these research variables. The targeted respondents of the questionnaires are R&D members in companies belonging to technology industry (electronic components and computer, communication, and optical products manufacture), traditional manufacture industry (mechanical equipments and metallic material manufacture), Internet and information services. In this study, we sent 300 copies of questionnaires by e-mail and post and received 210 returned questionnaires. Among them, 6 questionnaires were deleted due to ineffective or absent of their answers. This results in 204 effective questionnaires and a respondent rate of 68%.

RESULTS

Reliabilities

Before doing the hypotheses-test analyses, the reliabilities of research constructs were tested. The reliability of each construct was calculated using Cronbach's coefficient alpha. During the process of reliabilities test, we delete some of the items that are inconsistent with most other items and result in lower reliabilities of the focal variables in question. The remaining items are tested again, and the results are much better than the former.

The results are (1) interaction = 0.706, trust = 0.778, perspective taking = 0.556, with a uni-dimensional scale of social capital = 0.797, (2) observation = 0.767, conversation = 0.796, practicing = 0.833, bargaining = 0.800, with a uni-dimensional scale of learning activities of the CoPs = 0.892, (3) R&D performance = 0.894. From the results we find that the reliability of perspective taking of social capital is lower than 0.6 and indicates a moderate level of data quality, thus the analyses results that related to perspective taking should be interpreted with caution. All the other research constructs' Cronbach's alpha coefficients are greater than 0.7, ranging from 0.797 to 0.894, and comfortably demonstrate internal consistency of measurement.

Multiple Regression Analyses

The relationship between learning activities of the CoPs and social capital (H1)

Table 1 is the summary of the results of multiple regression models for H1. We note that apart from H1-4, in which all the three constructs of social capital (interaction, trust, and perspective taking) are significant related to dependent variable observation, the other three sub-hypotheses are only partially supported for their regression coefficients of the interaction construct are insignificant. In sum, among the three aspects of social capital, trust and perspective taking have positive impacts on

all four aspects of learning activities of CoPs, while interaction largely has no effects on learning activities of CoPs with only a mild effect on bargaining (p=0.025). Thus, the hypothesis H1 is partly supported.

Table 1 The results of multiple regression models for H1

	Model 1 (H1-1)	Model 2 (H1-2)	Model 3 (H1-3)	Model 4 (H1-4)	
Dependent variable Independent variable	Observation	Conversation	Practicing	Bargaining	
	(Beta)	(Beta)	(Beta)	(Beta)	(VIF)
Interaction	.114	.077	.122	.147*	1.146
Trust	.247**	.209**	.167*	.286**	1.369
Perspective taking	.193**	.212**	.161*	.228**	1.277
Adjusted R^2	0.173	0.142	0.104	0.249	
F	15.171**	12.165**	8.835**	24.433**	

N = 204 *$p<0.05$, ** $p<0.01$

The relationship between learning activities of the CoPs and R&D performance (H2)

The results of the regression model (dependent variable = performance) on the Table 2. It indicates that, among the four constructs of learning activities of the CoPs, both observation and bargaining positively affect R&D performance. This means the more observation and bargaining of learning activities among participants of an R&D project will result in higher R&D performance. And, "conversation" and "practicing" of learning activities of the CoPs have no significant impact on R&D performance. The possible reasons for failing to support H2-2 and H2-3 may be that conversation and practicing of learning activities of the CoPs need longer time to accumulate the knowledge and experiences of the focal professional area thus their effects may reveal in the future R&D projects not the current ones. Hence, H2 is only parly supported.

Table 2 The results of multiple regression models for H2

	Standardized Coefficients (Beta)	Collinearity Statistics (VIF)	
Observation	.227*	2.033	H2-1 (Supported)
Conversation	-.086	1.920	H2-2
Practicing	-.104	2.130	H2-3
Bargaining	.319**	1.685	H2-3 (Supported)

Dependent variable = R&D performance *$p<0.05$, ** $p<0.01$
N=204; Adjusted R^2 = 0.116; F=7.629 (0.000)**

CONCLUSIONS

Social capital will affect learning activities of the CoPs

In essence, three constructs of social capital will affect learning activities of the CoPs, while interaction just partially impact to learning activities of the CoPs, the other two (trust and perspective taking) constructs both significantly affect learning activities of the CoPs.

From the empirical results of this study, we can see a conspicuous fact that members of the CoPs where trust level high, people are more willing to engage in social exchange and interact with others, such as sharing the ideas, solving the difficult problems, having spontaneous conversations, and even sharing professional information and knowledge. In accordance with Mascitelli's (2000) study that indicates that more interaction among participants of an R&D project has a cumulative effect on prompting knowledge sharing, these conducts will positively impact learning activities. It will make members want to share professional knowledge and experiences by using a lot of ways to let the others such as apprentices to learn. And, because perspective taking may result in the behavioral effect of empathizing with other members and the cognitive effect of mutual understanding among members, it can facilitate the learning activities of the CoPs. Therefore, the more interaction, trust, and perspective taking of social capital will affect the more learning activities of the CoPs for gathering the professional knowledge.

Learning activities of the CoPs is positively related to R&D performance

The results confirm that some aspects of learning activities of the CoPs ("observation" and "bargaining") contribute to R&D performance. These results accord with the prevailing perspective that the members of the CoPs exhibit high level of learning activities can enhance the R&D performance.

The rejections of the association between "conversation" and "R&D performance" and between "practicing" and "R&D performance" may imply that conversation among members of the CoPs maybe just transfer some forms of information that have little effect on R&D performance. It may be the case that to promote R&D performance needs some deep mutual expertise and mutual understanding that can be generated from conversation to know more expertise, but conversation alone would hardly contribute to R&D performance. And the reason that practicing has no effect on R&D performance may be that to promote R&D performance needs longer time to let expertise keep practicing until they can internalize the mastery of that professional expertise. Thus, practicing may contribute to future R&D projects but hardly contributes to current ones' R&D performance.

*** References upon request**

The Exploration of Causes of the Performance of TQM Implementation In Hospitals In Taiwan

Shih-Wang Wu
Department of Hospital and Health Care Administration,
Chia Nan University of Pharmacy & Science, Taiwan
P.O.BOX 62-13, Tainan, Tainan City 70899, Taiwan
+886-6-2664911 ext.5225, scottwu101@gmail.com

Yafang Tsai
Department of Health Policy and Management,
Chung-Shan Medical University, Taiwan
No.82, Sec. 2, Chongsing Rd. Taichung City 404, Taiwan
+886-4-24730022 ext.17176, avon611@gmail.com

ABSTRACT

The administrators perceived the turbulent environment and difficult to operate after the implementation of NHI (National Health Insurance). The president of hospital usually adopted TQM (Total Quality Management) to improve the competence because that there is always many miscommunication, low efficient and waste besides the changing of environment. We try to explore the causes that can effect the performance of TQM implementation, and expect to suggest how to modify the way to adjust and plan. We adopt the structural questionnaire to collect the data, and try to understand the influence of TQM perception on the operation performance. We take staffs who have worked over six months in three regional hospitals that applying TQM over one year in middle Taiwan as our objectives. The significant results showed us that male, elder, higher education and married staff showed higher TQM perception. The staffs belong to paramedical department showed higher TQM perception than nursing department. Finally, we found the staffs' perception of TQM activities will positively influence the performance of TQM implementation significantly.

Keywords: Total Quality Management (TQM), Perception, Performance

INTRODUCTION

The health care service is a special kind of service industry in the customer-orientation era. The administrators of hospitals set the goal to satisfy their customers, because of customers' requirement are not only on medical technique and professional knowledge, but also on service quality. Implementing TQM (Total Quality Management) will lead to higher quality, and will give customers more guarantee and trust, and will positively influence the performance finally. Thus, the health care services can be improved wholly and continuously and pursue quality and performance by implementing TQM. It also can improve the satisfaction of internal and external customers. The study tries to explore if the difference of individual characteristics will show different degree of perception on TQM activities, and tries to explore if staffs' perception on TQM activities will influence the performance of TQM implementation.

LITERATURE REVIEW

TQM implementation in health care industry

Carey and Teeter (1995) implement TQM in Martin Router Hospital to reduce the mistakes successfully. They proposed the model of TQM provide a best process management of nursing quality. According to the findings of previous researches, we can understand that quality improving and TQM implementation is the best method to compete with others, and we can design the best way of operation to reach our goal.

The TQM Activities

The characteristics of health services are division of labor and team cooperation, TQM activity is, too. Stephanie et al. (2000) proposed the professional developing activities can push staffs to explore other skill, knowledge and great attitude besides their basic capability. They also emphasized that increasing the chance to develop staffs' professional skill in complicated medical environment will improve staffs' job satisfaction. The TQM activity is a kind of professional developing activities. The higher the perception of staffs on TQM activity is, the more the professional skill training. The staffs will show better job satisfaction, and will satisfy their customers.

The Performance of TQM

TQM as the resource of potential competitive advantage could face their customers better, improve customers' satisfaction, strengthen internal communication, and motivate

staffs' participation. It also could make staffs more aggressively, solve problem better, decrease cost, and make closer to their suppliers. All of these will improve organizational competitive advantage, and we also can adopt these to be indicators to evaluate our performance.

METHODOLOGY

We adopted structural questionnaire to collect our data. We modified the items according to professional suggestion and pretest of questionnaire, and accomplish the final version with Likert's five point scale. Cronbach's α of questionnaire are all higher than 0.7, they all show good reliability. We take staffs who have worked over six months in three regional hospitals that applying TQM over one year in middle Taiwan as our objectives. We issued 180 questionnaires and 140 valid questionnaires were collected with a response rate 77%.

We propose three hypotheses as follows.

H1: Different individual characteristics will show significantly different degree of perception on TQM activities.

H2: Different department characteristics will show significantly different degree of perception on TQM activities.

H3: The degree of staffs' perception on TQM activities will significantly influence the performance of TQM implementation.

RESULTS

The significant results of ANOVA on degree of perception on TQM activities showed that several different individual characteristics will show significantly different degree of perception. Male staffs are better than female; staffs' age higher than 41 years-old show higher degree; staffs graduated from university and institute show higher degree than whom educational degree is high school; married staffs show higher degree than single. The H1 is partially supported. There is just one significant result on department characteristics, which is the Para-medical department show higher degree on customization of service process than nursing department. The H2 is partially supported. The results show the degree of staffs' perception on TQM activities will positively influence the performance of TQM implementation significantly. The result of regression analysis is as table 1. The H3 is totally supported.

Table1. Regression Analysis of the degree of staffs' perception of TQM activities and the performance of TQM implementation

Independent variable	Dependent variable	R^2	Beta	P value
The degree of staffs' perception on TQM activities	The performance of TQM implementation	0.447	0.669	0.000*

*** P<0.005

DISCUSSION

The results showed that different individual characteristics will show significantly different degree of perception of TQM activities. We think that male are always more aggressive, so that they will show higher degree of perception than female. Staffs are older than 41 years-old show higher degree because of their rich experience. The staffs' education degree is higher than university show higher degree of perception of TQM activities. We believe that is because their well training of systematic improving, so they show higher degree. About the marriage status, the married staffs show higher degree than single. We think that because the married staff care about their job and hunger for stable work, so that that will concern their organization more than single.

The results showed that different department characteristics will show significantly different degree of perception of TQM activities. The Para-medical department show higher degree on customization of service process than nursing department. We believe the reason is nurses always do their work by physician's order, so that they can not adjust their work according to customer's opinion immediately. The staffs belong to Para-medical department can rearrange their process easier because of their independence. Maybe the administrators could improve the efficiency by rethink the process of departments' working.

The results show the degree of staffs' perception on TQM activities will positively influence the performance of TQM implementation significantly. We believe that staffs who understand the importance and benefit of TQM will show higher degree of perception on TQM activities, and will create better performance.

CONCLUSIONS

If the TQM be well controlled, the staffs will perform better. We recommend that if administrators wish their organizations could exist longer and get more competitive advantage, they should ask themselves how many additional value can be provided to their customers. The purpose of TQM is to improve the operation quality, increase the medical quality, meet their customers' need, and continuously innovate via staffs' effort.

Finally, we proposed three conclusions according to the discussion as follows.

1. The elder, married and well education staffs show better performance what the administrators want. The administrators should keep the excellent staffs, and adjust their work and position to strengthen their service quality. On the other hand, the administrators should create good culture and training system, and retain appropriate rotation rate in order to nurture more excellent staffs.
2. The administrators could take the para-medical departments to be benchmark of hospitals. Other departments could visit these departments and learn from them well, especially nursing department. The administrators also can rethink and redesign the process of nursing, and strengthen their integrated thinking in order to improve the whole medical service quality.
3. The higher the staffs' perception of TQM activities is, the higher the staffs' degree of adaptability and performance of TQM implementation. The administrators should announce appropriately and have a common consensus, thus will increase the efficiency and effectiveness of TQM implementation.

REFERENCE

Barney,J.B.(1986),Oraganizational culture: can it be a source of sustained competitive advantage, *Academy of Management Review*,11,656-665.

Cameron, K. S.(1985) Cultural Congruence Strength and Type Relationships to Effective, Working paper, Business School, University of Michigan.

Carey, R.G., and J.L. Teeters. 1995. CQI Case Study: Reducing Medication Errors. *Joint Commission Journal on Quality Improvement,* 21(5):232–7.

Chatman, J.A.and Jehn, K.A.(1994)Assessing the relationship between industry characteristics and organizational culture: how different can you be, Academy of Management Journal, 37,522-553.

Deal, T. E. and Kennedy, A. A.(1982),Corporate Cultures, Addison-Wesley, Mass。

Dension, D. R.(1984) Bringing corporate culture to the bottom line, Organizational Dynamics, 13, 5-22.

Gordon, G.(1985), The relationship of corporate culture to industry sector and corporate

performance, Gaining Control of the Corporate Culture, Jossey-Bass Publishers, San Francisco, 103-125.

Kotter,J.P.amdHeskett,J.L.(1992),Corporate Culture and Performance,Free Press,NewYork.

Nunnally,J.C.(1978). Psychometric Theory, 2nd ed.,New Your: McGraw-Hill.

Petrock, F.(1990), Corporate culture enhances profits, HR Magazine, 35, 64-66.

Quinn, R. E.(1988), Beyond Rational Management: Mastering the Paradoxes and Competing Demands of High Performance, Jossey-Bass Publishers, San Francisco.

Schein,V.E.(1985)Oraganizational realities:the politics of change, Training and Development Journal, 39, 37-41.

Scholz, C.(1987), Corporate culture and strategy – the problem of strategic fit, Long Range Planning, 20, 78-87.

Wallach, E. J.(1983), Individuals and organizations: the cultural match, Training and Development Journal, 37, 29-36.

MOMENTUM EFFECT OF LOTTERY-TYPE STOCK

Chun-An Li, Department of Finance, National Yunlin University of Science and Technology, Taiwan
Chia-Fei Lin, Department of Marketing and Logistics Management, Hisn Sheng College of Medical Care and Management, Taiwan
Che-Hung Hsu, Department of Business Administration, National Yunlin University of Science and Technology, Taiwan
Kun-Mei Pan, Department of Accounting Information, National Taichung Institute of Technology, Taiwan.

ABSTRACT

Based on Kumar's (2008) definition of lottery-type stocks: (1) idiosyncratic volatility, (2) idiosyncratic skewness, and (3) stock price, and Fama-French four-factory model and Harvey and Siddique's two-factory model, we identify Taiwan stock market securities into lottery and non-lottery two different types of stock. Then combine those stocks with mispricing or not further to four types of stock. The purpose of this paper is to investigate the four types of stocks which will perform momentum effect well in short-term or reverse in long-term. The empirical result shows that momentum strategy for lottery type stocks performance is more significant than non-lottery stocks, so take the zero-cost strategy to buy lottery type stocks which is undervalued and selling non-lottery type stocks which is overvalued can earn abnormal returns, especially holding this strategy returns is more significant in twelve month. The robust test results show the return of the strategy is affected by business cycle and January effect. Overall, our evidences shows that Taiwan stock market exist momentum effect in long-term undervalued.

Keywords: lottery-type stock、momentum effect、undervalued、overvalued

INTRODUCTION

Recently, lottery tickets are in vogue. They have low price opposite in the highest payoffs, thus people like to buy the lottery tickets. We employ measure of lottery stock following Kumar (2009): low stock price, high idiosyncratic return volatility, and high idiosyncratic return skewness. When the stock price has the bigger volatility expressed that the stock will possibly occur the more returns. The investor may use the small capital to obtain the higher payoff.

However the stock price has the high volatility that may assist mispricing effect. These stocks tend to be overvalued or undervalued and therefore available to arbitrage. Because the stock's

price derivate from its fundamental, we should discuss the momentum effect in the mispricing. Additionally, the characteristic of lottery stocks has the idiosyncratic volatility. With this motivation, we use the mispricing effect to inspect whether can cause lottery stock profits used by momentum effect.

Moreover, Weng et al. (1995) discussed that the price of the stock market is dividend into five groups; we found the lowest price in the annual return higher than the highest return group. The paper discussed this lottery-type stocks are among the low-priced. As mentioned above, we think that it should have the exciting results.

This study is based on stocks listed in Taiwan and dividend into two categories, the lottery-type stocks and non-lottery-type stocks. We combined with the mispricing effect on stock which is divided into overvalued and undervalued to explore lottery and non-lottery stocks will be higher. We think that we should buy lottery-type stocks to gain the more profits. The stock at the high volatility, there may exist a large investment opportunities. We can composition of the portfolio through the mispricing effect. And in the past studies have shown that low-priced stock returns are higher than the high-priced stocks. These factors are included in the lottery stock. In this study, we were should be able to use buying the undervalued of lottery-type stocks and selling overvalued of non-lottery stocks to get the higher returns. Finally, we also observe the environment impact in the stocks. The paper explore the momentum strategy is significant under the factors, multi-bear market and January effect.

DATA AND METHODOLOGY

To avoid the noise and price limit, the data used in this study are month observations covering the period from January 2005 to December 2009. We exclude certain unexpected events on the stock impact. And the study research the long-term abnormal returns used by month returns, not daily returns.

This reference to the Conrad and Kaul (1993) and Loughran and Ritter (1996), the paper modified the research framework, we dividend the stock market into the lottery and non-lottery portfolio. In addition, we added the book to market ratio (BM Ratio) to measure the stock is overvalued or undervalued. When BM Ratio is low, stock prices tend to overestimate the value for the fundamental price. Conversely, when the BM Ratio is high, the stock would tend to underestimate. This state makes four different portfolios, lottery undervalued group, lottery overvalued group, non-lottery undervalued group, non-lottery overvalued group. Each portfolio will be taken of all the stocks composite equivalent weighted and market capitalization-weighted. We observe the momentum strategy have significant profits in a year holding period. Finally, combine momentum strategy, we observe the business cycles and January effect will become factors to determine whether the sample at different times under the momentum strategies have significant rewards.

EMPIRICAL RESULTS

We use Jenter (2005) approach, adding the BM ratio as an indicator for justifying which stock is overvalued or undervalued. When the less of the ratio was overestimated, the greater the ratio was underestimated. Therefore, according to the shares of BM ratio for the classification criteria, arranged in descending group, 50% of the underestimation of the sub-group, 50% of overestimate of the after-group. Therefore, there are four kinds portfolio, lottery overvalued, lottery underestimated, non-lottery overestimated, underestimated non-lottery. And each portfolio of all stocks is grouped with equal weighted and value weighted, respectively, hold for one year, and calculate the buy and hold portfolio during the month cumulative return.

The study found that: regardless of whether there is over- or underestimate, lottery stock can be significant excess return relative to the non-lottery stock. In the lottery group were divided into overestimate and underestimate, we can be seen when the time stretched, lottery underestimate the group were also inferior to overestimate the performance to high-performance group, a difference also in the holding period of one year and 7%. By contrast, the value is non-weighted lottery method which can be watched to see whether a combination of high underestimated the inside, during the holding period returns are negative.

In this study, four groups on the use of the building momentum strategy are: buying lottery underestimate stocks and selling non-lottery underestimate group, buying lottery underestimate stocks and selling non-lottery overestimate group, buying lottery overestimate and selling non-lottery underestimate group, buying lottery overestimate stocks and selling non-lottery overestimated groups. We observe whether the constructed portfolio can still benefit significantly.

According to empirical results show that among the four groups of momentum strategy return per year can reach 25% or more, the effect is very significant. However, buying lottery underestimate stock and selling non-lottery overestimate is the highest of four groups. This momentum strategy in the four groups within a year is a long-term continued to rise, but there has been no reversal in the status kinetic energy. In general, we find the difference is not significant between equal weighted and value weight, probably because of the lottery-type stocks filtered out when the stock price is relatively low, so there is no difference regardless of portfolio return is get form equal weighted or value weighted.

For make sure the results are robust, we also test whether the results will change or not under different external environment affected. The empirical results show that:

(1) The January effect: comparison in January and non-January return on the investment portfolio, we can find whether in January and non-January with significant excess returns can be obtained.

(2) Bull market and Bear market comparison: We can find in the long period of any investment portfolio, the return was significantly beat the broader market.

SUMMARY AND CONCLUSIONS

In this study, to test the lottery and non-lottery stocks, mispricing effects, are combined into the momentum strategy performance, research period from in January 2005 to December 2009, shares of listed companies in Taiwan month return. We used the BM ratio to determine whether the stock market is underestimated or overestimate. And we test whether the groups was continued to reap excessive profits. Also in the final with a robust test to see if it will be subject to external environmental factors.

We can observe whether the value-weighted or equal-weighted equivalent, holding that it does not contain January and February, the momentum strategy is the same as the all period. If contains January and February, the returns significant higher than containing January and February. So we found if the matter with the whole investment period and the stock market has almost the same. But start investing in January is a relatively high return. Therefore, this study constructed the momentum strategy returns are the best to buy lottery underestimate and to sell non-lottery overestimate group at any time. The stock price of lottery underestimated investment holding period is significant reaction, instead of lottery overestimate group is overreacting. Finally, this study suggests investors can take advantage of buying a lottery underestimate and selling non-lottery overestimate portfolios. This group of strategy in Taiwan's stock market should be for the most significant profitability.

REFERENCES

1. Baker, M. and J. Wurgler, 2006, “Investor Sentiment and the Cross-section of Stock Returns,” *Journal of Finance* 61, 1645-1680.
2. Chan, L. K., N. Jegadeesh and J. Lakonishok, 1996, “Momentum Strategies,” *Journal of Finance* 51, 1681-1713
3. Conrad, J. and G. Kaul, 1993, “Long-Term Market Overreaction or Biases in Computed Returns?” *Journal of Finance* 48, 39-63.
4. Cutler, D., J. Poterba, and L. Summers, 1990, “Speculative Dynamics and the Role of Feedback Traders,” *American Economic Review* 80, 63-68.
5. Carhart, M. M, 1997, “On Persistence in Mutual Fund Performance,” *Journal of Finance* 52, 57-82.
6. Debon, W. F. and R. Thaler, 1985, “Does the Stock Market Overreact?” *Journal of Finance* 40, 793-808.
7. Dahlquist, M, and P. Sellin, 1994, “Seasonalities in Swedish Stock Returns: Why are they not

Arbitraged Away?" Seminar Paper, 583, Institute for International Economic Studies, University of Stockholm, Sweden.

8. Fama, E. F. and K. R. French, 1993, "Common Risk Factors in Returns on Stocks and Bonds," *Journal of Financial Economics* 33, 3–56.
9. Harvey, C. R., and A. Siddique, 2000, "Conditional Skewness in Asset Pricing Tests," *Journal of Finance* 55, 1263–1295
10. Ikenberry, D., J. Lakonishok and T. Vemaelen, 1995, "Market Underraction to Open Market Share Repurchase," *Journal of Financial Economics* 39, 181-208.
11. Jegadeesh, N. and S. Titman, 1993, "Returns to Buying Winners and Selling Losers: Implication for Market Efficiency," *Journal of Finance* 48, 65-91.
12. Jenter, D., 2005, "Market Timing and Managerial Portfolio Decisions," *Journal of Finance*, 60, 1903-1949.
13. Kumar, A. and C. M. C. Lee, 2006, "Retail Investor Sentiment and Return Comovements," *Journal of Finance* 61, 2451-2486.
14. Kumar, A, 2009, "Who Gambles in the Stock Market," *Journal of Finance* 64, 1889-1933.
15. Loughran, T. and J. R. Ritter, 2000, "Uniformly least powerful tests of market efficiency," *Journal of Financial Economics* 55, 361-389
16. Ritter, J., 1988, "The Buying and Selling Behavior of Individual Investors at the Turn of the Year," *Journal of Finance* 43, 701-717
17. Ritter, J., 1989, " Portfolio Rebalancing and the Turn of the Year Effect," *Journal of Finance* 44, 149 -166.
18. Rozeff, M., 1986, "Tax Loss Selling: Evidence from December Stock Returns and Share Shifts," In Proceedings of the Seminar on the Analysis of Security Prices. Center for Research in Security Prices, 9-45.

Influencing Factors of Shanzhai Products Perceive Values and Purchasing Intentions

Arthur Cheng-Hsui Chen, Department of Business Administration, National Yunlin University of Science and Technology
Shu-Fang Luo, Department of Business Administration, Tainan University of Technology
Yu-Hsiang Mao, Department of Business Administration, National Yunlin University of Science and Technology

ABSTRACT

With the financial crisis of 2008, a new consumption trend in Chinese culture called shanzhai products is threatening well-known brands in the market. Shanzhai product has the similar out-looking or designs of a well-known brand's product. The key difference versus counterfeits is shanzhai product does not have the brand logo, which creates a blurred line to judge if the shanzhai product is illegal. The objective of this study is to examine the different purchase intent to shanzhai products between individual and group buying situations, and the moderating effects by the logo/no logo and consumer moral beliefs. A mixed between and within experimental design, 4 (individual situation, group members with high recognition, low recognition, and without opinion) ×2 (with logo and without logo)×2 (high and low morality), was used to test the hypotheses. Undergraduate students were recruited as experimental subjects via convenient sampling, resulting to 595 valid samples. The results indicate that the purchase intention of group situation is significantly higher than that of individual situation. The higher recognition by the group members will enlarge their purchase intention vs. lower recognition group. The difference between individual and group purchase situation will be moderated significantly by the logo/no logo and consumer moral belief.

Key Word: Shanzhai products、counterfeits、knockoffs、consumer moral belief.

INTRODUCTION

In recent years, brand marketing in the emerging Asian markets has been catching manufacturers' attention. In particular, luxury brands are becoming more popular. However, with the financial crisis of 2008, a new trend called shanzhai products is threatening those well-known luxury brands in the market. Why these products are called shanzhai? In Chinese, shanzhai originally meant the mountain hiding place of bandits. But now this word is used to describe the products that copy the out-looking or designs of a well-known brand's product. In some way, these shanzhai products are very similar to that stealing some assets from a popular brand. Examples like shanzhai computers and shanzhai cell phones, which sometimes were

called “white brand” or “no brand”. In fact, consumers notice that these products, not being certified by the brand company, are of a lower quality and less reliability. However, consumers still like these products because of their low prices. Therefore, for those consumers who cannot afford to buy a popular brand due to their limited salaries, shanzhai products often become an alternative.

In the consumers’ mind, shanzhai products are thought as affordable innovative products, luxury alternatives. Without copying the brand name, this shows there is still a difference between shanzhai and counterfeit goods. Perhaps shanzhai is much similar to knockoffs in the Western culture. However, whether shanzhai products are legal or not is difficult to judge, due to different regulations in different countries. From a practical point of view, the market of shanzhai products does exist indeed. According to estimations, knockoffs product around the world is worth at least $9 billion (Wilson 2007); showing there is a research value in the consumer behaviors of these products. Therefore, legal issue will not be discussed in this research, but instead to explore the factors influencing consumers purchasing shanzhai products.

Many research have studied the attitudes and consumer buying behaviors toward counterfeit goods(Ang, Cheng, Lim, & Tambyah, 2001；Bloch, Bush, & Campbell,1993；Eisend and Schuchert-Guler, 2006；Tom, Garibaldi, Zeng, & Pilcher, 1998; Wee, Tan, & Cheek, 1995), but few are targeted on shanzhai and knockoffs. We believe that consumer behavior research with counterfeit goods may not be totally applicable to shanzhai products. This has triggered the motivation to study the consumer behavior of buying shanzhai products.

The most popular reason for consumer to buy counterfeit goods is hoping to be grouped as those who can afford luxury goods or to catch other’s attentions, even if they know they need to consider some moral or legal risks (Bloch, Bush and Campbell, 1993; Hoe, Hogg and Hart, 2003; Penz and Stotinger, 2005). However, someone buying shanzhai product is because the product have similar out-looking or design of a brand’s product, but with a much lower price. On the other hand, consumer is difficult to judge the legal marginal line for shanzhai product, how the impact does by their morality and the law is still unknown. Therefore, this study attempts to explore if the hesitation to buy shanzhai product will be more obvious during personal purchases, but increasing the possibility to buy for group situation due to share the risk consideration, which may cause the difference of purchase intent between personal and group situation. In addition, this study would like to further examine if the different degree of the member consensus toward the shanzhai product will affect their purchase behavior within a group. Thirdly, the shanzhai product will become the counterfeit product if they copy the logo from the well-known brand, which definitely deepening the moral and legal risks. Will this logo or no logo moderate the above relationship? Finally, in terms of consumer characteristics, will their level of moral beliefs affect their perception of legal risks and result in different behavior of purchasing shanzhai product? Therefore, this research will also examine if consumer moral belief will moderate the impact of the individual and group situation on the purchase intent to

the shanzhai product. In conclusion, the main purpose of this study is to examine the different purchase intent to shanzhai products between individual and group buying situations, and the moderating effects by the logo/no logo and consumer moral beliefs.

LITERATURE REVIEW AND HYPOTHESIS

Shanzhai product refers to products which have similar functions or designs with brand products. Being similar to brand products, they are often misunderstood as counterfeit. What makes shanzhai products different from counterfeits are that shanzhai products have been improved and innovated, and are more similar to a knockoffs.

Effects of Purchasing Situations

When consumers purchase shanzhai products, they may go by themselves or go with a group of 4-5 persons; what is the difference between the two situations? Consumers refer, rely, and listen to opinions of reference group when making purchasing decisions（Arndt,1967；Brown & Reingen, 1987；King & Haefner,1988；Kwong, Yau, Lee, Sin, & Tse, 2003；Murray,& John,1990）. Woodside & Delozier(1976) believe group decisions are far more willing to bear risks than individual decisions. Because risky shift and diffusion of responsibility often happen during a group decision, group members tend to be more willing to accept the riskier alternative (Kogan & Wallach, 1964). Group purchases also create more impulse buying and larger purchase expenses (Granbois, 1968); therefore a consumer in a group purchase situation should be willing to try a riskier shanzhai product and have higher chances of impulse buys than in an individual situation.

It is generally believed there are risks to bear when buying a shanzhai products, including psychological risk, social risk, financial risk, performance risk, physical risk , convenience risk, time risk and general comprehensive risk (Cox, 1967；Peter & Tarpey, 1975；Pires, Stanton & Eckford, 2004；Stone & Gronhaug, 1993). Buying a shanzhai product could have the financial risk where the price is not worth value of the product, the physical risk and performance risk when the shanzhai products performance and functions are not as good as the name brand, and most importantly the psychological risk related with pride and the social risk related to the perception of other-orientation. Consumers are still willing to purchase shanzhai products when they know there are risks involved because shanzhai products are usually trendy items, and young people like to compare with each other their new and interesting belongings. When the consumer does not have the shanzhai products his or her fellow peers have, there is a psychological risk of being mocked as out as date and untrendy. When peers recognize shanzhai products not as counterfeits but as affordable luxuries, shanzhai products gets popular within young people, and there is no social risk from criticism of buying shanzhai products. Group purchase situations can lower the social and psychological risk of buying shanzhai products and can raise the perceived value and purchase intentions.

Buying with others compared to buying alone seems to be less prone to mistakes. Or it can be said group purchase situations can diffuse the performance risk of buying poor quality shanzhai products, the financial risk of price not worth the product, and the risk of buying the wrong thing. With more people, everyone gets braver, even when a shanzhai product of lower innovation is bought, because everyone is buying it, the perceived value and purchase intentions are raised. But when an individual is purchasing, he alone has to bear all the purchase responsibility risks, but the perceived value and purchase intentions remain the same. Therefore, under a group purchase situation, there should be a higher level of perceived value and purchase intention. Hypothesis 1 is shown as follows:

H_1: During a group purchasing situation, the perceived value and purchase intentions will be higher than that of individual purchasing situations.

Group Recognition Effect

Individual perceived value and purchase intentions towards shanzhai products in a group purchasing situation should be highly related to the attitudes of the group members. Arndt (1978) believes the effects of a group are big mainly because group opinions are seen as trustable and reliable information that can help persons make better purchase decisions and provide more social supports. More social support also means more social pressure and surveillance. Group opinions provide more buying influences than personal belief (Bourne, 1957, Myers & Reynolds, 1967). The social role theory explains that an individual will be set and expected to have personal behaviors just like their role assigned. In order to maintain good interaction and communication in the group, the individual will perform the actions expected by group members, especially when one is follower or obedient, conformity behaviors show easily (Lippa, 1990).

When college students have more interaction with peers, and wish more to be accepted by peers, they are more likely to be obedient, and produce conformity behaviors (Lascu & Zinkhan, 1999). Peers with strong-tie produce group pressure situations on individuals because they have the need of belongings (Allen, 1965), and expect individual buy the same products and not to buy the different products with peers (Pincus & Waters, 1977). An individual scared of showing opposing opinion and tends to be compliant the group norms (Venkatesan, 1966). The support and recognition of peers toward a product affects the purchase intentions of an individual (Masrigal, 2000; Kelman, 1961); in other words, when ingroup peers have high recognition for shanzhai products, it is providing the sign that shanzhai products are reliable, and shows the group members expect individuals to follow and make purchases. And the individual, wishing to be identified by peers and seek a sense of belonging, will raise their perceived value and purchase intentions towards shanzhai products. When a ingroup peers have low recognition for shanzhai products, it is denying the value of shanzhai products and producing an opposing pressure of individuals purchasing shanzhai products. The individual, wishing to be identified by peers and seek a sense of belonging, will be afraid to show their opposing opinions and

follow the group's opinions and lower their purchase intention and perceived value toward shanzhai products. Therefore, hypothesis 2 is shown as follows:

H_2: When group members have high recognition towards shanzhai products, their perceived value and purchase intentions will be higher than that of low recognition of shanzhai products within group members.

The Moderating Effect of Logo

The main difference between shanzhai products and counterfeits is that counterfeits use the brand name of the genuine product while shanzhai product does not. But sometimes because of the extremely similar appearances of the shanzhai products with the genuine product, they are misunderstood and stigmatized. From the past literature shows that the main reason for consumers to buy counterfeits is that they value how brands can reflect individual status, but they cannot afford the genuine product (Albers-Miller, 1999 ; Bloch et al.,1993; Cordell et al.,1996, Kaikati ＆ LaGrace,1980; Wee et al., 1995). This shows they desire to be identified or fit into others group (Bloch, Bush & Campbell, 1993; Hoe, Hogg & Hart, 2003; Penz & Stotinger, 2005). As for the product, counterfeits are often low quality copies of genuine products with high brand value (Lai and Zaichkowsky, 1999), these brand name items are usually popular but hard to access, and has a respectable brand name (Rust, Zeithaml, & Lemon, 2004；Wan, Luk, Yau, Tse, Sin, Kwong & Chow, 2009). Sometimes, the shanzhai product also wishes to follow the trend of the popular brand, and "steal" its brand value. But at the same time it does not wish to bear the crime of counterfeit, so it will not put the logo on the shanzhai product.

As proposed in hypothesis 1, this study believes the purchase intentions towards shanzhai products differ greatly during group and individual purchasing situations. This is because shanzhai product is not necessarily illegal, it is just on the brink of the law system; the protection of a group will urge the individual to make riskier purchases. However, for counterfeits, because it bears the value of the brand name, but it also bears the risk of the law, it is much riskier than shanzhai product. Therefore, consumers will feel more protection in a group, with the risk bore by all members of the group, the consumer is more willing to go after counterfeits that have a higher brand name worth. In other words, towards counterfeits, consumers are attracted by the value of the brand name, and the group offers protection. We expect both combining effect will increases the difference of the perceived value and purchase intention of individual purchasing situation from group situation. As for shanzhai product, because there is no brand name involved, the consumer does not need to take as much risks and does not rely on group protection as much. Therefore the difference between group purchasing and individual purchasing situations will not be as large as that of counterfeits. Therefore, hypothesis 3 is shown as follows:

H_3: The presence of the authentic logo will affect the perceived value and purchase intentions of individual and group purchasing situations. In other words, the authentic logo will magnify the perceived value and purchase intentions in a group purchasing situation.

This study assumes (hypothesis 2) that from the view point of conformity behavior, the purchase of shanzhai product will be affected by the degree of consensus of the group members. However, since counterfeits bear the brand logo and increases legal risks, the affect of group approval could be different. From literature related to counterfeits, those who purchase counterfeits value their face and the view of others, so they wish to use counterfeits to gain face from others or fit into a group (Wen et al, 2009). However, courage is needed when purchasing counterfeits, because it is not easy to do illegal acts. Past researches have discovered significant effects of peers towards purchasing illegal goods (Albers-Miller, 1999). Therefore, when the group recognition is high, it not only adds courage, but also lowers the guilt of committing illegal acts; on the other hand, the attraction offered by the brand name magnifies the intention to purchase counterfeits. But if the group recognition is low, not only can not the risk be dispersed, the view of others and face consciousness can add to the consumer's concerns, and magnifies the consumer's original view point towards illegal acts (Tan, 2002), and lowers the intentions of purchase. Therefore, this study suggests the degree of identification within a group has greater effect on counterfeits than shanzhai products. That is the positive effects of a group with high approval and the negative effects of low approval will both be larger than shanzhai products, because the effects of legal risks and views of others is larger. Therefore, hypothesis 4 is shown as follows:

H_4: The presence of an authentic logo will affect the difference of perceived value and purchase intention of a group purchase situation with higher or lower levels of approval. That is, the authentic logo will magnify the perceived value and purchase intention of a group with higher degree of recognition.

Effect of Consumer Moral Beliefs

It is discovered from many studies relating to counterfeits that moral ethics, belonging to consumer's personal characteristics, is an important influential factor. Just as expected, the levels of consumers' moral ethics are related to the willingness to purchase counterfeits. There are studies that show the lower the ethical belief level of consumers is, the more likely is it for them to pirate software and purchase counterfeits (Ang et al., 2001; Kwong et al., 2003; Thing & Yap, 1998; Wagner & Snader, 2001; Wilcox et al., 2009). This study also wishes to discuss how ethical beliefs affect group and individual purchasing situations.

Muncy & Vitell (1992) defines consumer ethics as "acts of ethics principles and standards that lead consumers when they receive, use, and discard the products and services." However, the effects of ethics intensity is a multi-dimensions reconstruction (Jones, 1991), organizational

factors include group dynamics, authority factors, and social factors. In addition, individuals will be affected by the ethics beliefs and decide whether to follow the trends of the group, and groups will have strict cultural values which affect the individual's ethics attitudes and behaviors(Hofstede & Hofstede, 2005). Some studies even found that public opinion pressure of the group has greater effect on moral attitude (Swinyard, Rinne, & Kau, 1990) and ethics standards (Zey - Ferrell, Weaver, & Ferrell, 1979) than the individual's belief. In addition, the individual will be willing to change his belief and standards for the group he belongs in (Folkes & Kiesler, 1991). This shows that the different levels of ethics belief towards shanzhai products will have different effects on group and individual purchasing situations.

This study assumes group purchasing intentions will be higher than individual situations towards shanzhai products. But from the levels of ethics belief, the results may be different. For example, when ethics belief is high, even though the group has the ability to diffuse the risks; but on the other hand, high levels of ethics belief can urge the consumer to stay away from potentially illegal acts, causing group pressure on the individual to lower, and thus shrink the difference between group and individual purchasing situations. When ethics beliefs are low, not only will the tension of ethics be reduced, the group's risk dispersion ability will also grow, and therefore adding to the difference between group and individual purchasing actions. Therefore, hypothesis 5 is as follows:

H_5: When the consumer has low moral beliefs towards purchasing shanzhai products, the perceived value and purchasing intentions of a group purchasing situation will be higher than individual purchasing situation. When the consumer has higher ethics belief towards shanzhai products, the perceived value and purchasing intentions of a group purchasing situation will have no obvious difference from an individual purchasing situation.

In addition, from the perspective of group recognition, this study expects that when the consumer has a higher level of ethics belief, the tension of inner value will be larger, thus causing the attitude towards shanzhai products to become negative, and offsets the positive effects of high degree of group recognition and minimizing the difference of low degree of group recognition. Conversely, when consumers' ethics beliefs is low, not only is the inner tension reduced, the outer tension of group recognition increases, causing the purchase intentions of a group with high levels of approval to have a bigger difference than those with lower levels of group approval. Therefore, hypothesis 6 is as follows:

H_6: The levels of ethics of a consumer will affect the purchasing intentions of a group with higher or lower degrees of approval. When the consumer has lower ethics, the higher or lower level of approval within the group will have bigger effect on purchasing intentions. When the ethics of the consumer is high, the higher or lower level of approval within group members will have a smaller effect.

RESEARCH DESIGN

There are three experimental variables in this study. First is the individual and group purchase situation. The group situation is also divided into high group identification, low group identification, and no comment. The second is whether the product bears a logo. The third is the moral level of the consumer. This forms a 4(individual, high group identification, low group identification, no comment)×2(with logo, without logo) ×2(high moral belief, low moral belief) mixed experimental design. The first two variables are between subject designs. The moral belief variable is within subject design, and is grouped afterwards with the median of the average of subjects' moral asked items as a standard to divide subjects into high moral standard group and low moral standard group.

With phones being the first of shanzhai products, and many are indeed found in the market, this study will use the most common shanzhai phone, the iPhone, as the study subject. The shanzhai iPhone will be specially marked as being only $100 dollars (the real one costs about $750 dollars), and the Apple company logo will be on the phone. For group purchase situations, a group of four subjects will proceed to a shanzhai phone retail; for individual situation only one subject will proceed. Group identification level will be measured with subjects individually describing their identification level of shanzhai phones. The measurement items of Ang et al(2009) and Wilcox et al (2009) was used to measure the moral levels of subjects. The seven point semantic differential scale method was applied, which includes the following three questions: moral/unmoral, correct/incorrect, meets conscience/contrary to conscience (α=0.887). The perceived value was used seven point Likert scale to measure, adopting from Dodds et al. (1991) and Grewal and Monroe (1998). The measurement items included value exceeds price, worth buying, and economical buy. Purchase intention include "I will buy the product" and "there is a high chance of buying"(r=0.8890).

The color-printed questionnaire was used in this study. In the beginning, there was a text statement of buy scenarios under group or individual situations. Then the experiment showed the front and the back pictures of the cell phone, identifying the logo or no logo product for the subjects. After that, participants would answer the questions of manipulation check and the measurements of dependent variables. The final part of the questionnaire was the items of consumer moral beliefs. Undergraduate students were recruited as experimental subjects via convenient sampling, resulting to 595 valid samples, and randomly being assigned to different experimental groups.

RESULTS

Figure 1: Effects of Group Purchasing Situations and Group Recognition (for H1 and H2)

	group/individual situation				the level of group recognition				
	group n=147	individual n=148	t value	P-value	Higher n=153	no comment n=147	lower n=147	F value	P value
perceived value	4.16	3.90	2.14	0.03	4.72	4.16	3.55	37.92	0.00
purchase intention	3.29	2.75	3.58	0.00	3.86	3.29	2.65	30.04	0.00

Figure 2: The Moderating Effect of Logo under Group Purchasing Situations (for H3)

	With logo				Without logo			
N of sample	group n=72	individual n=73	t-value	P-value	group n=75	individual n=75	t-value	P-value
perceived value	4.28	4.07	1.40	0.16	4.03	3.73	1.66	0.09
purchase intention	3.74	2.90	3.82	0.00	2.86	2.61	1.32	0.19

Figure 3: The Moderating Effect of Logo under Group Identity (for H4)

	With logo					Without logo				
	Higher n=78	no comment n=72	Lower n=75	F value	P value	Higher n=75	no comment n=75	Lower n=72	F value	P value
perceived value	5.11	4.30	3.63	32.21	0.00	4.31	4.03	3.47	10.53	0.00
purchase intention	4.32	3.74	2.76	26.57	0.00	3.38	2.86	2.53	8.34	0.00

Figure 4: The Moderating Effect of Intensity of Consumer Ethics under Group Purchasing Situations (for H5)

	Lower moral beliefs n=103				Higher moral beliefs n=192			
	group n=63	individual n=40	t-value	P-value	group n=63	individual n=40	t-value	P-value
perceived value	4.43	3.95	2.53	0.01	3.94	3.88	0.44	0.66
purchase intention	3.75	3.03	3.07	0.00	2.95	2.65	1.58	0.12

Figure 5: The Moderating Effect of Intensity of Consumer Ethics under Group Recognition (for H6)

	Lower moral beliefs n=187					Higher moral belief n=260				
	Higher n=66	no comment n=63	Lower n=58	F value	P value	Higher n=87	no comment n=84	Lower n=89	F value	P value
perceived value	5.15	4.45	3.62	29.72	0.00	4.39	3.95	3.51	12.87	0.00
purchase intention	4.61	3.75	2.70	34.57	0.00	3.29	2.95	2.62	6.07	0.00

CONCLUSIONS AND IMPLICATIONS

The objective of this study is to examine the different purchase intention to shanzhai products between individual and group buying situations, and the moderating effects by the logo/no logo and consumer moral beliefs. The results indicate that the purchase intention of group situation is significantly higher than that of individual situation. In particular, the higher recognition by the group members will enlarge their purchase intention vs. lower recognition group. When the shanzhai product has brand logo, which become a counterfeits, the difference between individual and group will increase. The effects of different degree of member recognition will also be moderated by the logo or no logo. We also found that the degree of consumer moral belief is one of the key factors affecting the difference between individual and group situations when buying shanzhai products.

This research, applying social learning theory, face consciousness and other-orientation, advances our incipient but growing understanding of why consumers buy shanzhai products differently between individual and group situation. Specifically, we go beyond the legal issue but focusing on consumer response to the risk attitude toward the vague law violation of shanzhai products. The consumer in the group situation, especially for those getting recognition from group members, demonstrates their higher purchase intention not only to the shanzhai products but even further to the counterfeits products. In other words, consumer will need group protection more when the purchase risk is higher. However, we also found that higher consumer moral belief will offset the positive group effects on buying shanzhai products. This indicates that the internal drive of moral belief will neutralize the external force of group effects. The higher the moral beliefs, the lower the group effects on the shanzhai products purchase. We believe that these findings do have some theoretical contributions to the consumer behavior research and conformity theory.

Although shanzhai products will have some negative impact on the brand marketing, the market is existing. This implies that brand marketer cannot avoid the competition from shanzhai products, especially for those luxury brands. Therefore, the key is to find the best market segmentation among brand products, shanzhai products, and counterfeits. The ethical education can be used not only to strengthen individual consumer moral beliefs, which will reduce the market size of shanzhai products and counterfeits, but also reinforce an individual to play an important role in a group buying situation and reduce its purchase intention. On the other hand, brand marketers can introduce masstige brands, using similar designs or out-looking of well-known brand (a kind of shanzhai product but with new brand), to attract those middle level consumer who cannot afford for the luxury brands but want to show their closeness to high class segment. We think these suggestions may create an aggressive strategy and thinking to confront the competitions of shanzhai wave.

Reference upon request.

The Key Survival Factors of Micro-enterprise: An Exploratory Study of Retailing Industry

Sung-Po Chen, Professor, E-mail:sungpo@mail.nou.edu.tw
Department of Management & Information, National Open University, Taiwan
Chen-Yi Peng, Master of Business Administration, E-mail:cht@mail.nou.edu.tw
Department of Business Administration, National Taipei University, Taiwan

Abstract

Being different from former researches related to the issues of key survival factors and new value creation, this study is trying to propose a more practical and feasible point of view to new micro-enterprises. This paper integrates the key successful factors of new ventures from Song et al.(2008) and the resources, opportunities and team model from Timmons(1999) into the conceptual framework, and explores the key survival factors that micro-enterprises face in the process of starting and operating enterprises. We use in-depth interviewing method to collect data from entrepreneurs who established micro-enterprises in retailing industry. After data collecting and analyzing, this study provides the key survival factors during the process of starting and operating the micro-enterprises on year-based. This paper also offers suggestions for entrepreneurs in managing new micro-enterprises and in reducing the risk of failures.

To explore the key survival factors and their problem solutions that micro-enterprises might face, the author employed an in-depth interviewing approach by interviewing 16 micro-retailers. The results showed that the key survival factors in the early process of starting new ventures are: (1) Micro-enterprises are unfamiliar with the industrial environment which leads to their disability to explore the opportunities, to acquire information rapidly. (2) Micro-enterprises cannot manage their resources because of lacking entrepreneurial experiences (or capability). After the early stage of starting new ventures, the improvement of micro-enterprises' management by learning and the reduction of external environment's uncertainty will be the key survival factors for micro-enterprises to seek for available resources. This finding is consistent with Timmons' entrepreneurial theory. The author found that the exploration and selection of opportunities are vital before starting new ventures; the combination of entrepreneurial team is the key factor in the early stage of starting new ventures; and the demand of resources is the most important thing after the early stage of starting new ventures.

Keywords: Micro-enterprise, Resources, Opportunity, Entrepreneurial Team, Key Survival Factors

Introduction

This chapter can be divided into three sections. In the first section, we illustrate the research background and motive. The research theory and practice are clearly described in this part. In the second section, the research questions and purposes are put into discussion. In the third section, we explain the research process and clarify the framework arrangement for each chapter.

A. Research background and motive

According to "2010 White Paper on Small and Medium Enterprises in Taiwan" published by Ministry of Economic Affairs, there are 88,531 newly-started small and medium enterprises (SMEs) in 2009, which is 99.75% of all newly-started ventures (88,754). The selling value of newly-started SMEs is NT$157,600,000,000, which is 67.68% of all newly-started ventures and comes mainly from domestic sales. In recent six years, the percentage of newly-started SMEs of all SMEs was the lowest in 2008; the number of percentage began to increase in 2009. It shows that people in Taiwan are still active in developing newly-created enterprises.

Additionally, based on the research report executed by Chung-Hua Institution for Economic Research, authorized by the Directorate General of Budget, Accounting and Statistics (DGBAS) of Executive Yuan in 2009, domestic enterprises in Taiwan are mainly the micro-enterprises with staff members under 5 people. This kind of micro-enterprise is around 77.91% of all industries in Taiwan. This percentage seems to be higher than past few years due to the recent Financial Tsunami in 2008. This financial crisis resulted in a changed economic framework and an increasing unemployment rate. People who lost their jobs try to start up new ventures for survival. However, only 70% of newly-started ventures can last for 5 years, which shows that main challenges for micro-enterprises come not from the timing, but from the issue of how to make the sustainable operation. Therefore, to promote business survival rate, various kinds of consulting strategies have been provided to micro-entrepreneurs. Those strategies include the well-known "Youth Enterprise Loan" and "Youth Enterprise Incubation Program" held by National Youth Commission. After the SME Administration of Ministry of Economic Affairs promotes the "Micro-enterprise Loan," the Council of Labor Affairs of Executive Yuan follows these micro-enterprise consulting strategies and then promotes the "Micro-enterprise Phoenix Loan" and "Micro-enterprise Phoenix Consulting Plan" to assist micro-entrepreneurs to start up new ventures. Those strategies intend to

reduce the failure rate, to reach the basic living requirement of self-employment, and to cut down the unemployment rate.

In recent years, although government endeavors to build up a friendly industry environment for micro-entrepreneurs to survive and actively promotes the consulting strategies for micro-enterprises, most policies for solving problems are made by the mode of enterprise loan. However, the urgent key survival problems faced by micro-entrepreneurs have not yet been solved by planning any consulting strategies so that even if the micro-entrepreneurs apply loan, they still fail due to inability to well manage the capital. In this way, the survival rate for micro-enterprises cannot be increased and the effect of government consulting strategies falls short of expectations.

B. Research questions and purposes

Looking back on former researches of micro-enterprises, most of them focus on enterprise motive and enterprise management & activities, exploring key successful factors and new value creation (Kai-Ling Xu, 2006). The past researches focus less on key survival problems faced by micro-entrepreneurs in the enterprising process. As most micro-entrepreneurs start the new ventures by themselves, the target in the enterprising process is to expand the survival period. When they fail to reach the target of "creating new value," the entrepreneurs turn to focus on solving the present key survival problems. By exploring key survival factors that the micro-entrepreneurs meet when creating a new business, this study highlights the importance of enterprise theory and practice, intending to provide entrepreneurs or those who want to start up new ventures with expectations on key survival problems faced in the enterprising process and to offer government a reference for planning complete consulting strategies.

To know the operation modes of domestic micro-entrepreneurs, purposes of this research are listed as following:

1. Exploring key survival factors faced by micro-entrepreneurs in the enterprising process
2. Analyzing problems-solving methods on key survival issues faced by micro-entrepreneurs
3. Comparing the difference and similarity of problems-solving methods on key survival issues faced by micro-entrepreneurs, based on different operation modes and industries.

Literature Review

A. Entrepreneurial opportunities

The entrepreneurship described by Kirzner (1973) is "the alertness to the undiscovered opportunities." Gruen et al. (2006) thought opportunities can lead to desired outcomes in any situation so they are often defined as "environmental goodness for achieving aims." Because the above mentioned scholars thought in different perspectives, we can see that the former thought opportunities can be controlled by individuals in some parts and can be created subjectively; while, the latter thought opportunities supposedly cannot be controlled by individuals at all.

To start a new venture is a kind of action implying that one individual can utilize opportunities and is willing to take the uncertainties (McMullen & Shepherd, 2006). The source of the entrepreneurial opportunities can be affected by two important factors: 1) "the change of external environment," which is the emergence of new opportunities caused by the change of external factors, e.g., the new entrepreneurial opportunities brought by changes of the market, the technology, the politics, the regulations, the society, the population structure, and the industrial structure (Shane, 2005). 2) "individual factor of the entrepreneur," including the social network, the personality, the prior knowledge…etc., which means an individual discovers, evaluates, and utilizes opportunities in the process of enterprising (Shane, S. & Venkataraman, 2000).

In addition, Sarasvathy (2003) pointed out three types of entrepreneurial possibilities—the opportunity recognition, the opportunity discovery and the opportunity creation. "The opportunity recognition" means to notice there is possibility of creating a new business, to improve current relations of the supply and demand, and to further find new potentials of profits (Christensen, Madsen and Peterson, 1994). "The opportunity discovery" means to find solutions for current problems (Hsieh, Nickerson & Zenger 2007). "The opportunity creation" means entrepreneurs are supposed to have the ability to foresee so that they can create valuable market opportunities. Song (2008) et al., in a meta-analytic research on successful factors in starting up new ventures, suggested that opportunities are positively associated with enterprise achievements, showing that opportunity plays an important role in the process of enterprising.

B. Entrepreneurial resources

Chandler and Hanks (1994) thought that entrepreneurs are supposed to try their best in obtaining and disposing various resources so that the newly-started ventures will grow up rapidly and make the largest profits. Entrepreneurial resources, either visible or invisible assets, include finance,

entity, technology, human resources, society, and organization resources; types of the resources are changing according to different stages of enterprising. In the early stage of enterprising, the invisible resources (e.g., technology, organization…etc.) are more important than the visible ones (Lichtenstein and Brush, 2001). In addition, entrepreneurs have to set up the fundamental resources-managing ability and possess the organizational ability and central competitiveness by combining and utilizing resources; so that they can meet the needs for all enterprising stages.

According to the above literature analysis, this study takes Lichtenstein and Brush's perspective to define "entrepreneurial resources," and clarifies that it includes the visible and invisible assets of finance, entity, technology, human, society, and organization resources.

C. Entrepreneurs and their Team

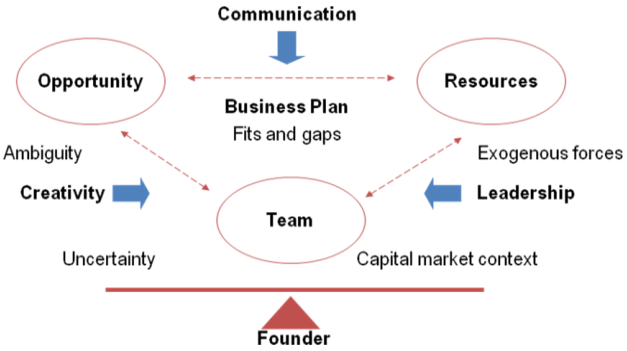
Bygrave thinks that entrepreneurs pursue opportunities and make profits unceasingly (1997). Concerning a "team," different scholars have different perspectives to define it. From a perspective of responsibility-partaking, they think that a team is a group of people working hard and sharing the responsibility of success and failure together for a specific objective and mission (McFletcher, 1996).

Salas (1992) proposed the characteristics perspective of team members and thought that a team is composed of at least two people with different backgrounds and characteristics. Katezenbach (1993), however, thought that a team is composed of a small number of people with complementary skills, who are the members recognizing the same objective and the process enabling them to share the responsibility with each other. This definition includes perspectives of "the same objective," "complementary skills," and "the responsibility-partaking," and is frequently adopted in the current literature on "team." This study adopts Katezenbach's perspective to define the "entrepreneurial team" as one composing of a small number of entrepreneurs who have complementary skills, pursue the same objective, work hard together, and take the outcomes collectively.

D. Resources-Opportunity-Team Theory

A famous enterprise management scholar in the US, Timmons (1999), mentioned in his book, "New Venture Creation," that the key to a successful enterprise lies in how the entrepreneurs grasp the three factors "opportunity, resources, and team" in the process of developing new ventures. Because effects caused by the environment and the time make changes on the newly-started ventures, the importance of three factors "opportunity, resources, and team" is different and imbalanced. With the development of enterprises, successful entrepreneurs have to make a flexible and a dynamically balanced arrangement on three factors "opportunity, resources, and team." In addition, changes never stop and keep occurring according to factors such as competitors and environments so the enterprising process is proceeding unceasingly.

According to the enterprise management model proposed by Timmons, it is considered most important to discover and choose the opportunity in the early stage of enterprising. In this process, the ambiguity of opportunity and the fast-changing of the environment make the enterprising process fully attacked and come with lots of risks. Thus, the success of a new venture relies on the leadership, communication, and creativity of the entrepreneurs themselves and also on their grasping opportunities and making good use of



resources, so that the business can proceed steadily. This study adopts Timmons' s perspective and proposes that entrepreneurs would develop a new venture by making a dynamically balanced arrangement on opportunity, resources and team.

Figure 2-1 Timmons's Model

E. Micro-enterprises

Definitions for micro-enterprises are different with countries. OECD defines a business with less than 20 staff members as a micro-enterprise. In Taiwan, only SME Administration of Ministry of Economic Affairs mentioned the term "micro-enterprise" in their 2003 promotion of "Micro-enterprise loan." It is defined in the SME Development Act IV, Paragraph 2, that manufacturing often employs staff members less than 20 people; agriculture, forestry, fishery, animal husbandry and service often employ staff members less than 5 people. Nowadays, we only define, in the SME Standard 3, the SME as one business with staff members less than 5 people, which is a definition

responding to the illustration on small-scale in the newly revised SME Development Act IV, Paragraph 2. The characteristics of micro-enterprise are low entry threshold and high survival difficulty and they face dilemmas of "small capital, weak constitution, and dispersive strength." This study adopts the definition used in the SME Act that the micro-enterprise refers to a business often with less than 5 staff members.

F. Key survival factors

Hui-Wen Lee (1994) takes service industry and Bing-Zhong Su (1995) takes manufacturing for example to analyze SME survival rate in Taiwan. Both of them adopt "Statistics on Sales Tax Expropriation Over Years" published by Financial Data Center of Ministry of Finance to count the survival period, based on the time interval of industry & commerce census (5-year unit). Yu-Zhen Shi (1996) studies the survival factors affecting SMEs in Taiwan and finds that factors of "education of decision makers," "whether to utilize finance resources issued by government," "whether to have functional offices abroad," "the enterprise finance in early stage," "age of entrepreneur," "debt rate," "whether to sponsor the public welfare," "the originality," "the entrepreneurship," "non-operating profits" have obvious influences on business survival ability. The "business survival" means one situation that after a certain period of operation the business still maintains the working status. Therefore, in the "2010 White Paper on SMEs in Taiwan," the number of SMEs is counted based on 2010 statistics of sales tax expropriation, provided by Financial Data Center of Ministry of Finance. In this study, indexes of performance evaluation are "whether the entrepreneurs can keep business running" and "whether the business can survive." Referring to the above-mentioned definition on business survival and using the statistics of sales tax expropriation over years, provided by Financial Data Center of Ministry of Finance, the research intends to understand "whether the business can survive" and compare it with business status defined by entrepreneurs themselves. This paper counts the survival rate by the certain time period, based on 1-year unit, and respectively discusses key survival factors for business operation of less than 1 year, 1~2 years, and 2~3 years.

G. Risks management

Williams et al. (1978) thinks that the risk is the "probability" of uncertainty that may cause damages to the person or industry in the future or the "undesirable outcomes" that may result in loses. In the "Risk Management：Concepts and Application" (1974) authored by two American scholars Robert I. Mehr and Bob A. Hedges, the definition is stated as that "the purpose of risks

management is to make economic assurance before the damages happen and to make the satisfied recovery after losses." Based on this definition, we conclude two purposes for risks management: 1) Pre-loss Objectives. 2) Post-loss Objectives. Risks evaluation is to expect the possible range for damages and estimate the potential losses to plan workable strategies to decrease undesirable influences. The Risk matrix combined by data of occurring frequency and range of losses is made to decide the degree of risks. The generalized classification of "the most serious," "serious," "moderate," "non-serious," "ignored" can be adopted to evaluate damages resulted from risks. The Risk matrix, with the combination of frequency and damage range, is usually used to evaluate risks. Based on the score value in the Risk matrix, it can be further sorted into several levels; and different management methods will be adopted for responding to different levels of risks. For example, The Critical Analysis (CA) usually applies Risk matrix to compare seriousness among all events. The comparison method is made by the relative concept, not the absolute concept. This relative concept is used for qualitative analysis because it is more flexible and easier to be put into practice and can be altered based on different situations.

Research Method

After explaining research purposes and reviewing the references, the research method is put into discussion in this chapter. The content includes building up the research framework, choosing the research targets and samples, and deciding the method for data collection.

A. Research framework

This study explores key survival factors for micro-enterprises by applying the qualitative research method, a method without statistics or any other quantitative process. Using qualitative research method, we can study life, behavior and interpersonal relationship of human beings. Once we use qualitative process to make analysis, whether the quantitative data are adopted or not, it can be called the qualitative research (Strauss & Corbin, trans. 1997).

This study intends to discuss key survival factors faced by micro-entrepreneurs in the process of starting up new ventures. The 24 key successful factors of newly-started ventures from Song et al. (2008) are integrated with "the resources, opportunity and team model" from Timmons (1999) to build up a conceptual framework and to explore key survival factors that micro-entrepreneurs face in the process of starting up and operating enterprises. The conceptual framework is presented as Figure 3-1.

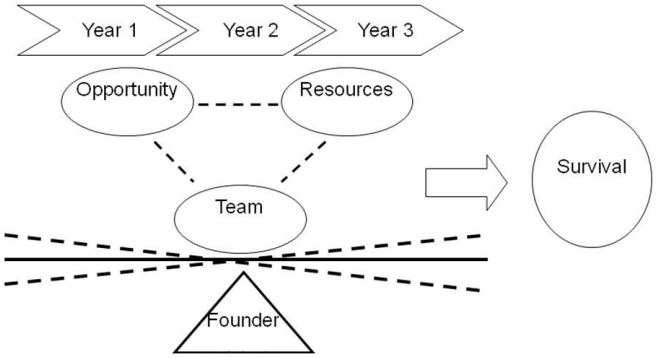


Figure 3-1 The conceptual framework

B. Research purpose

The research purpose of this study are entrepreneurs who establish micro-enterprises in retail industry. When micro-entrepreneurs join the "Micro-Enterprise Phoenix Program" held by Council of Labor Affairs of Executive Yuan, we ask for interviews with the micro-retailers and collect samples from those interviews to understand the detailed business running model and ideas. The period for interviewing is from Jun., 2010 to Aug., 2010. Sixteen micro-enterprises in retail accept interviews and can be the samples for this study.

C. Method for data collection

Shu-Man Pan (2003) thought that qualitative method for data collection is quite multi-dimensional. For social scientific studies, participation observation, in-depth interviewing, and documentary analysis are three main methods for obtaining research data. "Participation observation" means that

researchers, in a natural situation, make a systematical observing and recording process on the phenomena and behaviors. "In-depth interviewing" refers to an oral description method for interviews with particular targets to obtain a complete understanding. In the interviewing process, researchers would create a natural situation for interviewees to feel respected and equal so that the two-way interaction and communication can be made. "Documentary analysis" implies a research method applying any present data or documentary data to study social scientific issues. The so called "documentary data" covers a large range, including any present script data or any non-script data or records.

This study adopts the in-depth interviewing method and collects data by distributing the semi-structured questionnaires to explore key survival problems and decisions made by micro-entrepreneurs. The survey focuses on key survival factors of each year from the very beginning of the first-year enterprising. We make items for questionnaires based on yearly period fixed for enterprising and apply the open-answering model to collect data. The detailed items for questionnaires are as Appendix 1.

Conclusion

This chapter presents the analysis conclusion after organizing the collected data. The content includes the basis analysis on samples, the reliability and validity, and the risks management model.

A. Basis analysis on samples

The basis analysis on samples provides an integral overview on the sample structure. In this chapter, the population statistics variables include 9 items-- "gender," "age," "education," "entrepreneurial experiences," "history," "number of staff," "capital in record," "enterprising style," "vice-enterprise." Target research objects are sixteen micro-retailers, including costume industry, food industry, cosmetics industry, pet appliances industry, and etc. The capital of most interviewed micro-enterprises is below 1 million (87.5%; 1 million not included). The history of more than half of micro-enterprises is 1-2 years (62.5%). The staff of those micro-retailers is mostly the self-supporting entrepreneurs (43.75%). Only five micro-enterprises are consisted of 1 staff member without counting the entrepreneur himself/herself (31.25%). Most of the sixteen micro-retailers do not have any entrepreneurial experiences (75%). The detailed information is organized as Table 4-1. Refer to Appendix 2 for the micro-enterprise information collected from in-depth interviewing.

Table 4-1 Basis analysis on samples (n=16)

Gender of entrepreneur	Percentage
Male	0%
Female	100%
Total	100.0%

Age	Percentage
< 20-year-old	0%
≥ 20-year-old < 30-year-old	18.75%
≥ 30-year-old < 40-year-old	43.75%
≥ 40-year-old < 50-year-old	37.50%
≥ 50-year-old < 60-year-old	0%
≥ 60-year-old	0%
Total	100.0%

Education	Percentage
Junior High	0%
Senior High	56.25%
College	31.25%
Graduate	12.5%
Total	100.0%

Entrepreneurial Experiences	Percentage
Yes	25%
No	75%
Total	100.0%

Entrepreneurial Style	Percentage
Substantial Store	75%
On-line Store	25%
Total	100.0%

History	Percentage
< 1 year	25%
≥ 1 year to < 2 years	62.5%
≥ 2 years to < 3 years	6.25%
≥ 3 years to < 4 years	0%
≥ 4 years to < 5 years	0%
≥ 5 years	6.25%
Total	100.0%

Number of staff member	Percentage
0	43.75%
1	31.25%
2	12.50%
3	12.50%
4	0%
Total	100.0%

Capital in record	Percentage
≤ 0.5 million	43.75%
≥ 0.5 million to < 1 million	43.75%
≥ 1 million to < 3 million	12.50%
≥ 3 million to < 5 million	0%
≥ 5 million	0%
Total	100.0%

Vice enterprise	Percentage
Costume Industry	50%
Food Industry	25%
Cosmetics Industry	18.75%
Pet Appliances Industry	6.25%
Total	100.0%

B. Method for data analysis

Three methods for data analysis are adopted in this section, first including the content analysis method, and then the reliability and validity analysis, and the final one of risks management model analysis. We further explain the above-mentioned research methods in the following.

1. Content analysis method

This study systematically makes organization and analysis on the complicated interviewing information based on two stages of content analysis method. The first stage is the item analysis. We conclude properties from Timmons' s (1999) three factors "resources," "opportunity," and "team model" and then generalize items based on properties from 24 key successful factors of new ventures from Song et al. (2008). In the second stage of coding, we code based on items generalized from the first stage. In the end, we obtain 4 items of team property, 9 items of opportunity property, 7 items of resources property; there are 20 items in total. The definitions on the property and item are listed as Table 4-2.

Table 4-2 Code table for enterpreneurial models of property and item

Team		
No.	Item	Definition
ETM1	Industry experiences	Industry and marketing experiences of a company
ETM2	Marketing experiences	Marketing experiences of a company
ETM3	Entrepreneurial experiences	Previous entrepreneurial experience of a company
ETM4	Research & Development experiences	Research & development experiences of a company
Opportunity		
No	Item	Definition
EOY1	Competitive density	The competitive density among companies of an industry
EOY2	Environmental movement	The rapid environmental movement outside a company
EOY3	Environmental difference	The diversity and complexity outside a company
EOY4	Internationalization	The involvement of a company into transnational activities
EOY5	Low-cost strategy	Low cost as the competition superiority of a company
EOY6	Market increasing rate	The operating profits-increasing degree of a company
EOY7	Market range	The diversity and range of clients, client groups and company items
EOY8	Marketing density	The level of a company pursuing unique marketing strategies
EOY9	Product innovation	The level of a company developing new ventures and introducing new products and services
Resources		
No.	Item	Definition
ERE1	Financial resources	Assets inside a company
ERE2	Financial support	Financial support provided from business institutions
ERE3	Patent protection	Patents and technologies possessed by a company
ERE4	Strategy alliance	The cooperation and alliance of a company
ERE5	Investment on research & development	The density of investment in internal research & development activities of a company
ERE6	Supply chains organization	The cooperation between a company and upstream firms and between a company and downstream firms.
ERE7	Academia-Industry cooperation	The academia-industry cooperation of a company

After interviewing, this study organized an article about the interviews on the sixteen micro-retailers and then further made the coding. Two consultants with enterprises-consulting experiences make coding for this research. They receive professional training on related knowledge of resources, opportunity, and team and make communication on the coding regulations so that the consultants can completely understand the definitions of items, the codes, and the generalization of items. Therefore, consultants make judgments correctly. They code problems faced by the micro-retailers with the negative definition and code resolutions with the positive definition. After they finished coding, they make comparison with each other. If the two consultants have different comments on the coding, they will make discussion to reach a final consensus.

2. Reliability and validity analysis

Reliability refers to the estimation inaccuracy. When the inaccuracy numeric is smaller, the reliability becomes higher (Fang-Ming Hwang, 2007); this means that the estimation result consists of unity and stability. Validity refers to the estimation characteristics and behavior correctness. Even, the validity refers to the statistics index of whether exams can estimate the potential characteristics (Fang-Ming Hwang, 200). The reliability and validity should be put into consideration together for ensuring the examination quality.

This study adopts the interjudge reliability to make examination on the reliability. The interjudge agreeableness with item classifications between coding consultants is the standard accepted by the interjudge reliability. Refer to Table 3 for the calculation result. The interjudge agreeableness between coding consultants is higher than 90%, the reliability is above 95%, which reaches Kassarjian’s（1997）85% satisfaction standard.

Table 4-3 Table of reliability

Opportunity, Resources, Team
Interjudge agreeableness= (2X44)/(47+47)=0.93
Reliability= (2X0.93)/1+(1X0.93)=0.96

This study integrates “the resources, opportunity and team model” from Timmons (1999) as the property items and the key successful factors of new ventures from Song et al. (2008) as the classification method into a conceptual framework. This research framework is contended to possess validity.

3. Analysis on risks management model

This study adopts “Risk Matrix” to consider risks resulted from the combination of different “frequency” and “seriousness.” We calculate the frequency of key successful factors

mentioned by the micro-retailers and classify weights of key successful factors into four levels--level 5, level 3, level 1 and level 0 (the largest number refers to the most serious one). Based on the weight statistics and Timmons framework, we first make frequency & weight distribution figure and then further make the weighted "frequency" * "weight" figure. To make analysis, the "seriousness (weight)" as the Y-axis and the "frequency" as the X-axis are used to make risks management matrix distribution figure. The "." shown in the figure refers to the 0 (frequency) in the frequency figure, the 0 (seriousness) in the weight figure, the 0 (the weighted product) in the "frequency" * "weight" figure, the proportion location in the risks management matrix distribution figure. Other numbers respectively refer to frequency, weight, and the weighted product. Sizes of circles shown in the frequency figure, weight figure, and frequency" * "weight" figure are adjusted and standardized according to numbers.

4. Result of data analysis

a. The first year

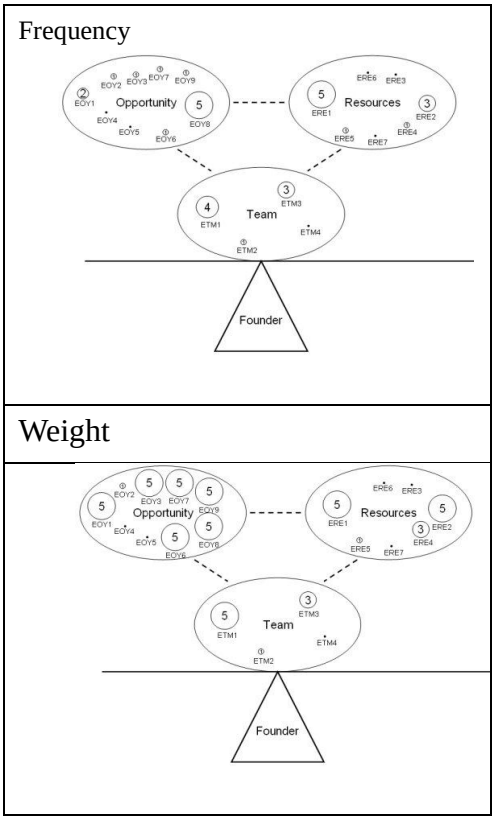


Figure 4-1 Key survival problems in the first year of enterprising (frequency and weight)

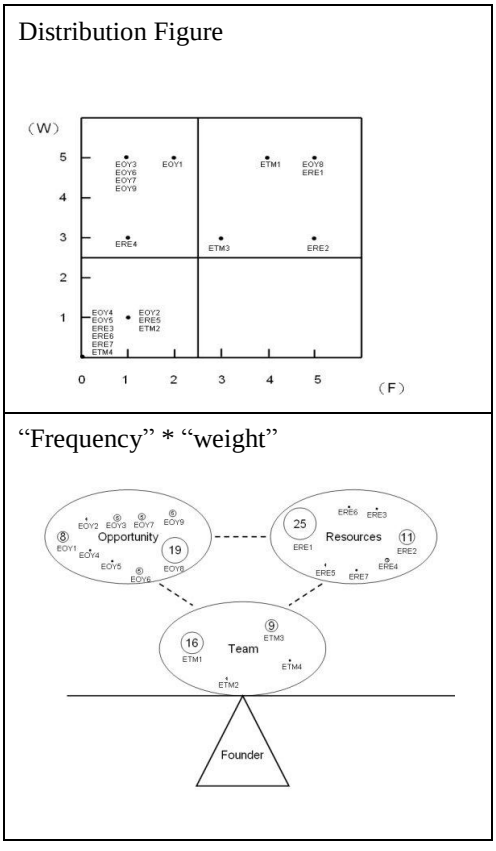


Figure 4-2 Key survival problems in the first year of enterprising (risks management matrix distribution—frequency*weight)

Refer to Table 4-4 for analysis on key survival factors for the first year of micro-enterprises in retail industry.

Table 4-4 Analysis on key survival factors for the first year of micro-enterprises in retail industry

Property	Item	Definition
Team	ETM1	Lack of team-managing knowledge of a company and not familiar with market situation
	ETM3	The team's lack of entrepreneurial experiences in a company
Opportunity	EOY8	Lack of particular marketing strategies of a company
Resources	ERE1	Unable to effectively manage and distribute the long term and short term capital
	ERE2	Lack of ability to collect capital (like gaining finance or making investments)

b. The second year

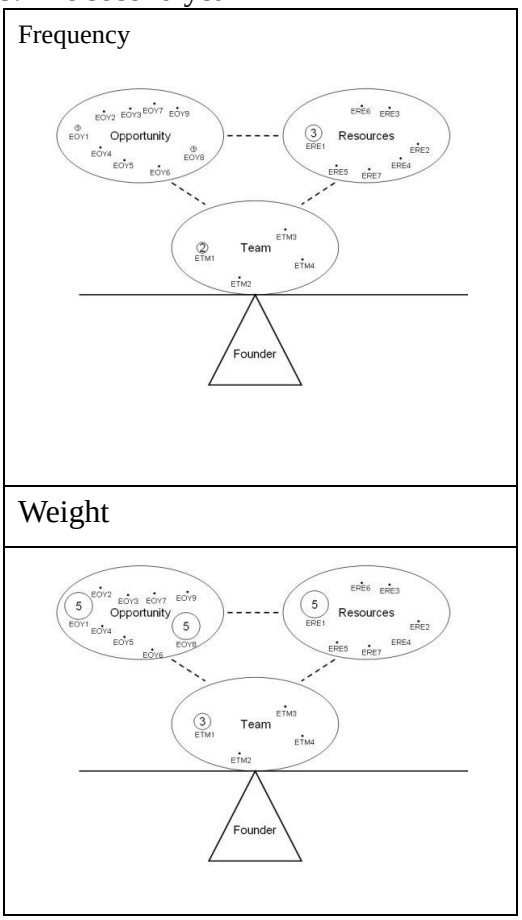
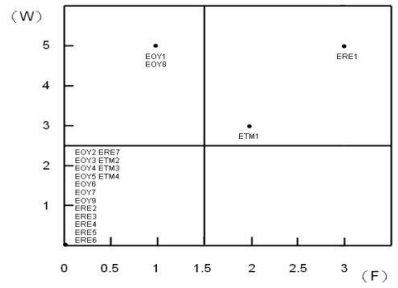


Figure 4-3 Key survival problems in the second year of enterprising (frequency and weight)

Distribution Figure



frequency * weight

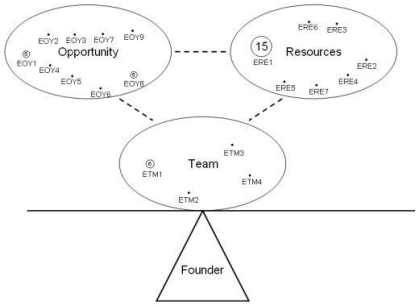


Table 4-4 Key survival problems in the second year of enterprising (risks management matrix distribution—frequency*weight)

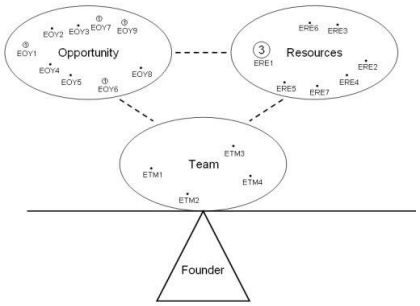
Refer to Table 4-5 for the analysis on key survival factors for the second year of micro-enterprises in retail industry.

Table 4-5 Analysis on key survival factors for the second year of micro-enterprises in retail industry

Property	Item	Definition
Team	ETM1	Lack of team-managing knowledge of a company and not familiar with market situation
Resources	ERE1	Unable to effectively manage and distribute the long term and short term capital

c. The third year

Frequency



Weight

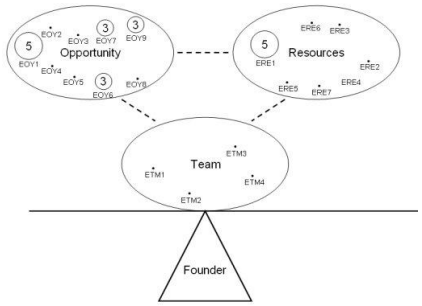


Figure 4-5 Key survival problems in the third year of enterprising (frequency and weight)

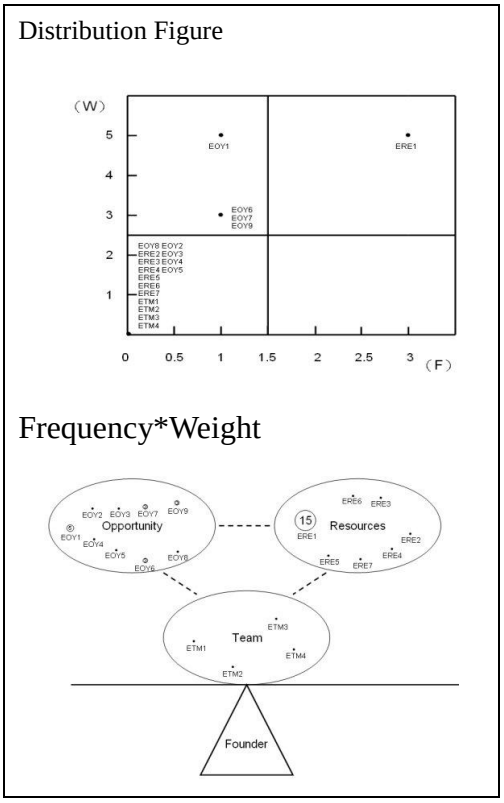


Figure 4-6 Key survival problems in the third year of enterprising (risks management matrix distribution—frequency*weight)

Refer to Table 4-6 for the analysis on key survival factors for the third year of micro-enterprises in retail industry.

Table 4-6 Analysis on key survival factors for the third year of micro-enterprises in retail industry

Property	Item	Definition
Resources	ERE1	Unable to effectively use and distribute the long term and short term capital

5. Conclusion and Discussion

After organizing data of key survival problems in the first year of enterprising, we find, in the "opportunity" property, that micro-entrepreneurs usually lack particular marketing strategies to seize opportunities (EOY8). In the "resources" property, micro-entrepreneurs are not good at distributing and managing capital (ERE1) and do not know how to find financial support (ERE2). In the "team" property, micro-entrepreneurs are not familiar with industry knowledge and market situation (ETM1); they even don't know how to deal with enterprising problems due to lack of related experiences (ETM3). Methods for micro-enterprising retailers to solve questions are to make particular marketing strategies to seize opportunities (EOY8) or to carefully increase income and decrease expenditure (ERE1) for survival. In the first year of starting up new ventures, most retailers are not well aware of the degree of industry competitiveness (EOY1).

Therefore, in the first year of micro-enterprise operation, key survival factors include "understanding degree of industry competitiveness (EOY1)," "making particular marketing strategies (EOY8)," "effectively distributing and managing capital (ERE1)," "whether to have financial support (ERE2)," "the familiarity with

industry knowledge and market (ETM1)," "whether to have entrepreneurial experience to deal with problems(ETM3)."

After organizing data of key survival problems in the second year of enterprising, we find, in the "resources" property, micro-entrepreneurs still are not good at distributing and managing capital (ERE1) but know how to find financial support. In the "team" property, although micro-entrepreneurs are not familiar with industry knowledge and market situation (ETM1), they know how to solve the problems. In the "opportunity" property, micro-entrepreneurs do not well understand the degree of industry competitiveness (EOY1) and do not have particular marketing strategies to seize opportunities (EOY8). However, problems faced are less than the first year. Methods for micro-entrepreneurs to solve problems are mostly by planning particular marketing strategies to seize opportunities (EOY8) or by finding capital to stably run business (ERE2).

Therefore, in the second year of micro-entrepreneurs operation, key survival factors include "effective distribution and management of capital (ERE1)", "familiarity with industry knowledge and market situation (ETM1)."

After organizing data of key survival problems in the third year of retail enterprising, we find that in the "resources" property, micro-entrepreneurs still are not good at distributing and managing capital (ERE1). In the "opportunity" property, micro-entrepreneurs can adapt to uncertainties caused from environments (EOY1). Although entrepreneurs may face those challenges, most of them will not become key survival factors. Compared with "opportunity" and "resources" property, the "team" property does not contain any key factors that will affect ventures' survival. Methods for micro-entrepreneurs to solve problems are mostly by planning particular marketing strategies to seize opportunities (EOY8) or to increase income and decrease expenditure (ERE1) for passing crises.

Therefore, in the third year of micro-enterprise operation, the key survival factor is "the effective distribution and management on capital (ERE1)."

6. Management Concept

This study helps micro-entrepreneurs foresee the key survival problems before and in the enterprising process to effectively manage the new ventures and increase the survival possibility. It also provides suggestions to governmental institutions to plan effective consulting strategies or to build up an environment suitable for micro-enterprises. The important concepts for new venture management practice are as following:

a. For micro-entrepreneurs

After organizing key problems faced by micro-entrepreneurs in the all stages of starting up new ventures, we conclude the research results. In the beginning stage of enterprising, micro-enterprises cannot well adapt to industry environment. For example, they cannot effectively seize opportunities, cannot catch opportunities rapidly, do not have sufficient enterprising experiences (or abilities), and cannot well use resources (cannot effectively manage capital). Those become key survival factors in the beginning stage. After the early stage of starting up new ventures, micro-entrepreneurs improve the inner organization by learning and intend to decreases influences caused from uncertainties of outside environments. In this stage, whether micro-entrepreneurs can seek suitable resources becomes key factor for survival. This proves Timmons' s enterprising theory framework arguing that the exploration and selection of opportunities are vital before starting up new ventures. The combination of entrepreneurial team is the key factor in the early stage of enterprising. After a business has been started and is in the operation process, the demand for resources becomes the most important issue.

This study also finds that it is the entrepreneur himself/herself that manages a new venture, even without an entrepreneurial team. The entrepreneurs improve their abilities in the process of operating a new venture to make up shortages caused from the lack of an entrepreneurial team.

In the early stage, micro-entrepreneurs face problems of marketing and channel management. Those problems resulted from micro-entrepreneurs' unfamiliarity with opportunities so that they highly and wrongly estimate the sales volume. Then, they would meet products-selling problems because of lacking workable channels. The financial management ability is one factor easily to be ignored by entrepreneurs; however, in the process of enterprising, it is definitely one important factor producing crisis and also being the turning point for business transformation.

Therefore, micro-entrepreneurs should develop industry knowledge and ability to build up a stable source of marketing information and then decrease risks of enterprising.

b. For government' s industry consulting strategies

In past few years, government has had many professional and practical consulting courses for new venture entrepreneurs. Those courses help a lot for entrepreneurs. However, those courses mainly focus on how to write enterprising proposal, but less pay attention to skills that entrepreneurs should have when starting up new businesses. They hardly remind entrepreneurs to pay attention to risks management in the process of making

entrepreneurial strategies. We suggest that consulting courses should help build up a concept of industry analysis before enterprising and make judgment on self-ability to avoid the rapid failure caused from blind investments. This study allows government to plan various consulting strategies for various stages based on different business operation modes, industry classifications, and the survival time. In this way, new micro-enterprises can be developed successfully, and the resources searching and analyzing abilities can be strengthened.

7. Research direction in the future

The below-mentioned three points are suggestions for the following researchers:

a. Our research targets are mostly female entrepreneurs joining the "Micro-Enterprise Phoenix Consulting Plan." The following researchers can consider enrolling micro-entrepreneurs other than members of the said plan so that the study will be more complete.
b. The following researchers can analyze and compare key survival factors by studying micro-entrepreneurs from different industries. Therefore, the similarity and difference of key survival factors of micro-entrepreneurs from various industries can be fully understood.
c. This study makes analysis on key survival factors by the view of "survival time." We suggest the following researchers can make analysis on key survival factors by adopting different points of view.

Reference

1. Strauss, A. & Corbin, J. (1990/1997). Basics of Qualitative Research : Grounded Theory Procedures and Techniques.
2. Brush, C. G., Greene, P. G. and Hart, M. M., (2001), From Idea to Unique Advantage: The Entrepreneurial Challenge a Resource Base, Academy of Management Executive, Vol.15 (1), 64-78.
3. Bygrave, W. D., (1997), The Portable MBA in Entrepreneurship, New York, NY John Wiley and Sons, Inc..
4. Bygrave, W. D. and Hofer, C.W., (1991), Theorizing about Entrepreneurship, Entrepreneurship Theory and Practice, Vol. 16 (1), 13-22.
5. Chandler, G. N. and Hanks, S. H., (1994), Market Attractiveness, Resource-Based Capabilities, Venture Strategies, and Venture Performance, Journal of Business Venturing, Vol.9 (4), 331-349.
6. Christensen, Peder, O. Madsen, and Rein Peterson (1994), "Conceptualizing Entrepreneurial Opportunity Identification", In Marketing and Entrepreneurship: Research Ideas and Opportunities, Gerald E. Hills, Ed. London: Quorum Books.
7. Cooper, A. C. and Arts, K. W., (1995), Determinants of satisfaction for entrepreneurs, Journal of Business Venturing, Vol.10, 439-457.
8. Gruen, T., Osmonbekov, T., and Czaplewski, A. E-wom, (2006), The impact of customer-to-customer online know-how exchange on customer value and loyalty, Journal of Business Research, 59 (4):449–456.
9.
10. Hsieh, C., Nickerson, J.A., and Zenger, T.R. 2007. "Opportunity discovery, problem solving, and the theory of the entrepreneurial firm", Journal of Management Studies, 44(7): 1255-1277.
11. Kassarjian, Harold H. (1977), Content Analysis in Consumer Research, Journal of Consumer Research, 4 (June), 8-18.
12. Katezenbach, J. R. , & Smith, D. K., (1993), The Wisdom of Teams: Creating the High-Performance Organization, New York: Harper Collins.
13. Lichtenstein, B. M. and Brush, C. G., 2001, How do Resource Bundles Develop and Change in New Ventures? A Dynamic Model and Longitudinal Exploration, Entrepreneurship Theory and Practice, Vol.26 (3), 37-58.
14. McFletcher, D., (1996), Teaming by Design: Real Teams for Real People in Chicago, IL: Irwin Professional Publishing.
15. McMullen, J.S. and Shepherd, D.A. 2006. "Entrepreneurial Action and the Role of Uncertainty in the Theory of the Entrepreneur", Academy of Management Review, 31(1): 132-152.
16. Murphy, G. B., Trailer, J.W., and. Hill, R.C., (1996), "Measuring Performance in Entrepreneurship Research", Journal of Business Research, 36 (1), 15-23.
17. Robert I. Mehr, Bob A. Hedges., (1974), Risk Management: Concepts and Applications, McGraw-Hill Inc.
18. Salas, E. et al., (1992), Toward an Understanding of Team Performance and Training, in R. W. Swezey and E. Salas eds. Team: Their Training and Performance, Norwood, NJ: Blex Publishing Corporation, 3-29.
19. Sarasvathy, S. D., Dew, N., Velamuri, S. R. and Venkataraman, S., 2003. "Three Views of Entrepreneurial Opportunity", in Acs, Z. J. and

Audretsch, D. B. (eds.), Handbook of Entrepreneurship Research, 141-160.
20. Shane, S. A., 2005, Finding Fertile Ground: Identifying Extraordinary Opportunities for New Venture, Wharton School Publishing.
21. Shane, S. & Venkataraman, S., (2000), The promise of entrepreneurship as a field of research, Academy of Management Review, 25 (1):217-226.
22. Song, M., Podoynitsyna, K., Hans van der Bij, & Halman, J., (2008), Success factors in new ventures: A meta-analysis, The Journal of Product Innovation Management, 25:7-27.
23. Timmons, J. A., (1999), New Venture Creation, 4th ed., Burr Ridge, IL：Irwin.
24. Venkatraman, N. and Ramanujam, V., (1986), "Measurement of business performance in strategy research: A comparison of approach", Academy of Management Review, Vol.11 (4), 801-814.
25. Wiklund, J. and Shepherd, D., (2003), Knowledge-based resources, entrepreneurial orientation, and the performance of small and medium-sized businesses, Strategic Management Journal, 24(13):1307－1314.
26. Williams J D, Reeves D S, Condie A P & Brumfitt W (1978), Infections of the Urinary Tract. Ed. E H Kass & W Brumfitt, University of Chicago Press; 8-18.

LOTTERY STOCKS IN THE TAIWAN STOCK MARKET

Chin-Shui Lo, Department of Finance, Chang Jung Christian University, Taiwan(R.O.C.), No.396, Sec. 1, Changrong Rd., Gueiren Dist., Tainan City 71101, Taiwan (R.O.C.), locs@mail.cjcu.edu.tw, Tel: 886-6- 2785-123
Chiung-Hui Cheng, Graduate School Business and Operations Management, Chang Jung Christian University, Taiwan(R.O.C.). No.396, Sec. 1, Changrong Rd., Gueiren Dist., Tainan City 71101, Taiwan (R.O.C.), Tel: 886-6- 2785-123

ABSTRACT

Many compare the stock market to gambling. Investors and gamblers may share common traits. This research investigates lottery stock behavior in the Taiwan Stock Market and finds market capitalizations and earnings to be good predictors of lottery stocks. Gambling-induced co-movements are significantly stronger for stocks traded by individuals than those traded by institutions. Consistent with the findings of [21], our empirical results suggest that individuals tend to gamble during an economic downturn. However, lottery stocks were found to only weakly exacerbate market volatility during up market periods, whereas non-lottery stock acts as a stabilizer for the market during an up period.

Keywords: Lottery stocks, gamble, volatility, skewness

1 Introduction

Some have compared the stock market to gambling, suggesting many commonalities between gambling and stock investment. The survey of [17] revealed that gamblers and investors share five traits and differ on three traits.[1] [26] advanced the commonality between lottery players and stock traders from the perspectives of overoptimism, aspiration, emotion, and entertainment. Many investors possess a risk-seeking propensity and wager on an uncertain investment outcome. The gamble of the remote chance of a large gain may be more attractive and expected for investors by taking a large chance of a small loss. The gamble habit might be deep rooted in the human psyche ([21]), be it a tradition of gambling in the Chinese New Year ([7]), a religious background with a gambling propensity ([23]), or a tendency to break-even ([28]). The trading volume of the Taiwan Stock Market has been ranked twelfth worldwide ([1]) and share turnover is over three times that of the developed market ([16]), such that high trading records might be due to investor preference of gambling-related stocks. This study investigates lottery stock behavior, which captures gambling characteristics and sheds light on their effect in the Taiwan Stock Market.

[11] proposed that people who buy insurance policies simultaneously buy lottery tickets, and are both risk averse and risk taking. Based on the insurance-lottery framework, [22] posits that, whether foolish or wise, people do participate in lotteries, not only insurance.

Prospect theory posits that value function describes why people have a risk-averse attitude for gains and a risk-seeking mentality for losses.

Stock trading is a negative-sum game ([26]). [11] advanced that people trade stock or buy lottery tickets because their trades provide the likelihood of promoting from the working class social position to the upper class. Investors may be reluctant to realize their losses ([25]). [19] reported that people gamble the game that they do not accept when they are in the gain domain, but want a ticket out of loss when they are in the loss domain. Gambling is a get-rich-quick game that attracts peoples to play. [2] conjectured that investors might overweight the low probability event and prefer stocks with positive skewness.

A lottery ticket represents a marginal cost that renders a remote chance of earning an extreme return ([6]). A parallelism can be advanced between investing in certain types of stocks and purchasing a lottery ticket. [17] indicated three requirements of gambling: gambling often depends on luck, the stake can be lost entirely, and a lucky bet can result in an excess return. However, in the stock market, investor gambling preferences for well-defined stocks as a lottery instrument is unobserved or unavailable. No specific stock is likely to possess the particular characteristics of lottery tickets, particularly the disproportional reward relative to cost. For these reasons, it is difficult to investigate the gambling-induced investment decision. In this vein, to indirectly identify stocks that could be perceived as a gambling object, [21] proposed three typical characteristics: low stock price, high idiosyncratic skewness, and high idiosyncratic volatility.

The evidence provided by [8] for low stock price reveal that investors in low–priced stocks yield lower (even negative) average returns as the premium paid for the chance to earn an extreme return. Compared to high-priced stocks, low-priced stocks are more attractive to investors with gambling preference, who are searching for cheap bets. The findings of [6] show that investors in the Taiwan Stock Market are willing to sacrifice average return for a remote chance to earn an abnormal return when they trade low-priced stocks. Based on these papers, low-priced stock serves to prime candidates of gambling-type stocks, which are more attractive to gamble-preference investors.

Based on prospect theory, [2] proposed that investors may have a preference for stocks with positive skewness in the return distribution. The model of [4] suggests that investor prefer skewness as an outcome of over-optimism about the probability of good states. To address this issue, [14] provided evidences showing that skewness of prize distributions explains why risk-averse individuals may play the lottery.

[15] suggested that because the variance and skewness of bet returns are highly correlated, bettors might appear to prefer variance when it is skewness that they covet. When stock has high volatility, investors might believe in a greater likelihood to realize previous extreme return events again, even though the re-occurrence of that event is likely assigned a considerably lower probability ([21]), while high return volatility likely emerges the dream of an extremely large payoff ([7]). Within the set of low-priced stocks, investors are more likely to select stocks with high idiosyncratic or stock-specific volatility. Among stocks that have low prices, high idiosyncratic skewness, and high volatility are more likely to be perceived as lotteries.

Risk aversion is the primary premise in finance theory. However, an understanding of investment behavior is incomplete when risk seeking attitudes rooted in the minds of investors are dismissed ([24). Research is increasingly concerned with the effect of speculative stock market activities. The Taiwan Stock Market has been viewed as one of the most volatile and speculative developing markets. During the Lunar New Year, Taiwanese always gamble because gambling and winning means good luck for the whole year. An interesting finding by [13] revealed that buying lottery tickets with a large jackpot can substitute for stock trading in the Taiwan Stock Market.[2] [1] found that the introduction of legal lottery in Taiwan reduced the TSE turnover by approximately one-fourth. However, empirical evidence of the lottery stock in this market is scant. This research investigates the existence and effect of lottery stock on return and volatility on the Taiwan Stock Market. Because of the lottery-type stock variance, there is approximately a 30% probability for a lottery stock to move forward to another type of stock. The evidence indicates that market capitalizations and earnings per share are good predicators of lottery-type stocks. Gambling-induced co-movements are also significantly stronger with lottery stock traded by individuals than that of institutions. Consistent with the findings of [21], our empirical results suggest that individuals tend to gamble during an economic downturn. Our findings also show that lottery stocks could weakly drive market volatility during up market periods, whereas the non-lottery stock acts as a stabilizer on the market during an up market period.

The remainder of this research is organized as follows. The next section presents the data and measures of lottery stock. Section 3 describes the findings of this research, and the final section concludes this research.

2 Data And Lottery Index

This research focuses on lottery stock behavior in the Taiwan Stock Market, an emerging market, motivated by the study of [21]. This work obtained daily trading data extracted from the Taiwan Economic Journal database (TEJ). The data set includes common stock listed on the Taiwan Stock Exchange Corporation from October 1986 to March 2010. From this period, the number of listed firms increased from 130 in 1986 to 763 in 2011.

To measure the index of lottery stocks, daily returns are first needed to estimate the idiosyncratic volatility and skewness. The quarterly volatility and skewness are the second and third moment of residual returns obtained by fitting a four-factor model to the daily stock returns. The four factors are market premium, size premium (small minus big), book-to-market premium (high minus low), and momentum (winners minus losers).

The quarterly idiosyncratic skewness and volatility are estimated based on the model of [3]. The model is defined as follows:

$$\mathrm{iv}_{i,t} = \left(\frac{1}{N(t)}\sum_{d\epsilon S(t)} \varepsilon_{i,d}^2\right)^{1/2} \quad (1)$$

$$\mathrm{is}_{i,t} = \frac{1}{N(t)}\frac{\sum_{d\epsilon S(t)} \varepsilon_{i,d}^3}{\mathrm{iv}_{i,t}^3} \quad (2)$$

where $iv_{i,t}$ and $is_{i,t}$ are idiosyncratic volatility and skewness, respectively. *S(t)* denotes the set of trading days from the first day through the end of quarter t, and let $N(t)$

denote the number of trading days in quarter t. Let $\varepsilon_{i,d}$ be the regression residual returns using four factors, which include the Fama and French three-factor model augmented with a momentum factor MOM.

Consistent with the method of [22] [23], this work assigns all stocks of listed firms each quarter to vigintiles (semi-deciles) based on stock price, idiosyncratic volatility, and skewness (where 1 is the lowest volatility and skewness groups, and the largest price group). The price, volatility, and skewness vigintile assignments are added quarterly for each stock to create a lottery score ranging from 3 to 60, which is then scaled to range from 0 to 1 applying (score − 3)/(60 − 3) (abbreviated as lottery index in following). A higher lottery index posits a stock that is more attractive to investors as an object to gamble. This study identifies stocks ranked by the lottery index in the lowest 30th percentile as non-lottery stocks, and the highest 70th percentile as lottery stocks. Stocks that do not belong to lottery and non-lottery stocks are defined as other stocks.[3]

3 Empirical Results

3.1 Basic Statistics

Table 1 reports the summary statistics for all variables,[4] the yearly average for all firms, and the average time-series over the sample period, examined in this research. Lottery stocks are defined as having a higher average lottery index, lower price, higher volatility, and skewness. On the contrary, the non-lottery stocks have a lower lottery index, higher price, and lower volatility and skewness. On average, stocks with higher turnover, liability, and book-to-market ratio are classified as lottery stocks. Table 1 also presents non-lottery stocks as having higher market capitalizations and earnings per share. Sales growth and returns do not significantly differ between lottery and non-lottery stocks, and are highest in the category of other stocks.

Table 1 Basic statistics

This table reports the mean quarterly characteristics of lottery-type stocks during the third season 1986 to the first season 2011. We identify stocks ranked by the lottery index as the lowest 30th percentile as non-lottery stocks, and the highest 70th percentile as lottery stocks. Stocks that do not belong to lottery and non-lottery stocks are defined as other stocks. The index is the lottery index. Std is idiosyncratic volatility. Skew is idiosyncratic skewness. Price is stock price. We assign all stocks of listed firms each quarter to vigintiles (semi-deciles) by stock price, idiosyncratic volatility, and skewness (where one is the lowest volatility and skewness groups, and the largest price group). Value is the natural log of market capitalization. Turn is share turnover. Lib is liability divided by asset. Eps is earnings per share. Sale is sales growth. Ret is stock return. BV is book-to marker ratio. P-value denotes the difference between lottery stock and non-lottery stock.

	Index	std	skew	price	size	turn	lib	EPS	sale	ret	bv
Non-lottery	0.297	6.643	6.462	6.836	9.276	23.389	38.852	1.869	15.884	-0.056	0.500
Other lottery	0.494	10.427	10.882	9.833	8.832	30.986	39.303	1.109	55.923	-0.034	0.627
Lottery stock	0.688	13.798	13.339	15.088	8.293	33.563	43.351	0.247	23.651	-0.067	0.947
p_value	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.153	0.766	0.000

3.2 Transition Probability Of Lottery Stocks

Lottery stocks are defined to capture three characteristics of small price, high idiosyncratic volatility, and high skewness. Because of the variations of the three gambling-related characteristics, lottery stock candidates likely vary over the period. Table 2 indicates the transition probability of three category stocks ranked by the lottery index. If a stock is identified as a lottery stock in quarter $t-1$, there is 60% probability that the stock is also a lottery stock in quarter t. The probability of the lottery stock moving to another stock is approximately 30%, but the probability for a stock to "jump" from a lottery stock to a non-lottery is the lowest of approximately 8%. The opposite is found in non-lottery stocks. For a non-lottery stock in quarter $t-1$, there are 10% and 38% probabilities to move forward from a lottery stock to another stock in quarter t. A 45% probability exists for another stock in quarter $t-1$ to stay in the same category as other stocks in quarter t. The next section evidences the transition probability of three category lottery stocks and examines what characteristics drive the transition.

Table 2 Transition probability of lottery stock

This table reports the transition of three categories of lottery stock during the third season 1986 to the first season 2011. Lottery, other, and non-lottery are lottery stock, other stock and non-lottery stock, respectively.

	Lottery (t+1)	Other (t+1)	Non-lottery (t+1)
Lottery (t)	0.600	0.318	0.082
Other (t)	0.254	0.455	0.291
Non-lottery (t)	0.103	0.378	0.519

3.3 Predictors Of Lottery Stocks

What features contribute to a stock being a candidate for a lottery stock? Table 1 shows different characteristics between lottery and non-lottery stocks; however, the characteristics of stock-level fundamentals and market factors might be a predictor of the lottery stock. To determine which characteristics are stronger predictors of lottery stocks, this work applies the [10] cross-sectional regression estimates to test what features might predict lottery stocks. We examine the prediction using the following regression model:

$$LOTs_{i,t} = Std_{i,t-1} + Skew_{i,t-1} + \mathrm{Pr}ice_{i,t-1} + MValue_{i,t-1} + Turn_{i,t-1} + \mathrm{Re}turn_{i,t-1} + BM_{i,t-1} + Lib_{i,t-1} + Eps_{i,t-1} + Saleg_{i,t-1} \quad (3)$$

The dependent variable is the index of lottery stocks (*LOTs*) for stock i in quarter t. The independent variables in the regression specification in quarter $t-1$include (1) three gambling-related characteristics: volatility (*Std*), skewness (*Skew*), and price (*Price*); (2) four market variables: natural log of firm market capitalization (*MValue*), share turnover (*Turn*), stock return (*Return*), and book-to-market ratio (*BM*); and (3) three firm fundamentals: liability-to-asset ratio (*Lib*), earnings per share (*Eps*), and sales growth (*Saleg*). Both the dependent variable and the independent variables have been standardized (the mean is set to zero and the standard deviation is one).

First, we run the quarterly cross-sectional regression and then compute the mean value of the coefficient and test the significance. Table 3 shows the empirical results. For the 98 quarters, we find that 93 coefficients on the price variable out of 98 quarters significantly and

positively associate with lottery stocks. The higher the stock price index, the lower the stock price. The magnitude of the mean coefficient on price is positive and significant at the 1% level. The evidence is consistent with the definition of lottery stock, and low-priced stock serves as a good candidate of lottery stock. The result is consistent with the findings of [21].

The 79 coefficients out of 98 coefficients on volatility are positive and significant at the 5% level. The mean of coefficients on volatility is significant at the 1% level and their sign is expected, indicating that volatility also serves as a good characteristic to identify a lottery stock. Among the three characteristics of lottery stocks, only seven of 98 quarters for the coefficient estimates for skewness significantly and positively differ from zero at the 5% level and their mean magnitude is much smaller than the other two counterparts. The overall magnitude of the coefficient on price is the largest among the three characteristics of lottery stock, and the small price is the cheapest wager to risk the gamble with a small probability of large gains.

Our regression results suggest that, for the coefficient estimates of market capitalization, 11 and 32 quarters of 98 quarters are significantly positive and negative, respectively. The mean of coefficients on market capitalization is negative and statistically significant, suggesting that small-cap stocks are more likely to become lottery stock in the future. From evidence provided in Table 3, only small-cap stocks tend to be a future candidate for lottery stock. The other three market variables have no capability to predict lottery stock. Among the three firm fundaments, a stock with high liability implying high risk is more likely to serve as a future candidate of lottery stock. Findings show a negative mean coefficient for earnings per share and their absolute magnitude is approximately triple that of liability, suggesting that a firm with high earnings is less likely to be a lottery stock in the next quarter.

Table 3 Prediction of lottery stock

This table reports the prediction of lottery stock during the third season 1986 to the first season 2011. LOTs is the lottery index. The remaining variables are the same as Table 1. Mean is the average of the time-series coefficient estimates. Positive (negative) is the number of positive (negative) coefficients estimated quarterly. $LOTs_t = Std_{t-1} + Skew_{t-1} + \mathrm{Pr}ice_{t-1} + MValue_{t-1} + Turn_{t-1} + Lib_{t-1} + Eps_{t-1} + Saleg_{t-1} + \mathrm{Re}turn_{t-1} + BM_{t-1}$

	Std	Skew	Price	size	turn	Lib	Eps	Sale	Ret	bv
mean	0.259	0.016	0.557	-0.04	0.013	0.019	-0.051	-0.006	0.003	-0.01
positive	79	7	93	11	10	15	0	6	11	3
negative	0	3	0	32	11	4	23	3	4	8
p-value	0	0.027	0	0.006	0.246	0.012	0	0.384	0.704	0.35

Because both lottery stocks and their explainable variables have been standardized, it is easy to compare the influence of each independent variable on the lottery stock. In summary, among the three firm fundamentals, the strongest statistical predictor for lottery stocks is earnings per share. The lower earnings a stock has, the more likely it is to be a lottery stock. However, market capitalization is the best proxy of market variables to predict lottery stock.

3.4 Determinants Of Return Co-Movements In Lottery Stocks

This study is motivated by [21], who investigated whether investor gambling activities

induce excess co-movements in stock returns. This subsection examines the determinants of return co-movements with lottery stocks. We first compute the quarterly lottery stock return co-movement measures for each stock i by estimating the following regression:

$$\mathrm{Re}t_{i,t} - R_{ft} = \alpha_0 + \alpha_1 RP_{i,t} + \alpha_2 SHM_{i,t} + \alpha_3 HML_{i,t} + \alpha_4 MOM_{i,t} + \alpha_5 LOTret_{i,t} + \varepsilon_{i,t} \quad (4)$$

The dependent variable is a daily risk premium for stock i in quarter t. In addition to the four factor model of [5], the independent variables include the daily portfolio returns of lottery stocks, which are identified at the end of quarter $t-1$. The main variable of interest α_5 from this regression is the return co-movement measures for stock i related to the portfolio returns of lottery stocks. We estimate the beta α_5 quarterly using daily returns where the return series used in the regression are standardized so that the co-movement estimates related to lottery stock return indices can be compared with other independent variables ([22]).

The findings of [22] suggest that return co-movements are stronger among stocks that have lottery features and that are located in regions where Catholicism is more prevalent. Individual investors accounted for over 60% of the trading volume, whereas margin trades are only allowed for domestic individuals in Taiwanese market. This study uses margin purchase traders as individual representatives. To explore the preference of individual investors toward lottery stock, we employ the following regression specification to examine the return co-movements of lottery stock related to the chosen variable. The dependent variable, beta, comes from α_5 in Eq. (4). *LOTs* denotes the index of lottery stock. *Ins* (*Ind*) is the holdings of institutional investors (individual margin traders). Remainders in the independent variables are the same of that in regression (3). We run the following regression from the first season of 1997 to the first season of 2011, because the available data sets containing holdings of the institutional investors begins from January 1997.

$$\begin{aligned} Beta_{it} = {} & \beta_0 LOTs_{it} + \beta_1 MValue_{it} + \beta_2 Turn_{it} + \beta_3 Lib_{it} + \beta_4 Eps_{it} \\ & + \beta_5 Saleg_{it} + \beta_6 \mathrm{Re}turn_{it} + \beta_7 BM_{it} + \beta_8 Ins_{,t} + \beta_9 Ind_{it} \\ & + \beta_{10} Saleg_{it} + \beta_{11} LOTs_{it} * Ins_{it} + \beta_{12} LOTs_{it} * Ind_{it} \end{aligned} \quad (5)$$

For too many people, gambling is a simple form of entertainment. Most evidence suggests that individual investors prefer to gamble, whereas recent evidence indicates that gambling attitudes also influence institutional investors ([22]). We apply regression (5) in the stock-level to examine return co-movements with lottery stock. In the regression, the dependent variable is the co-movement measure derived from regression (4) and we control for the direct effect of lottery index and individual holdings (institutions), and firm fundamentals and market variables in independent variables. Consistent with the foregoing regression, we apply the [10] cross-sectional regression estimates to examine what features might commove with lottery stocks. Both the dependent variable and the independent variables are also standardized. We are concerned with the interaction terms (β_{12}) of lottery stock (*LOTs*) and individual holdings (*Ind*); if individuals trade more lottery stock, we expect β_{12} to be significantly positive.

Table 4 presents the results. The coefficient of the *LOTs*×*Ind* interaction term in the

regression is highly significant and positive, indicating that co-movements are strongest for lottery stocks traded by margin purchase traders. However, observations also show that the coefficient (β_{11}) on the interaction term ($LOTs \times Ins$) of lottery stock and institutions is also highly significant but negative, suggesting that institutional investors do not exhibit a greater appetite for lottery stocks. The absolute magnitude of the joint effect of lottery stock and individual trading is larger than that of lottery stock and institutional trading, implying that the effect of a gambling-induced stock on excess co-movements is significantly stronger on individuals than institutions. In summary, individuals exhibit a greater appetite for lottery payoffs than institutions, a result consistent with the findings of [22], who documented that return co-movements are stronger among stocks that have lottery features and that are located in regions where people are more prone to gamble. Contrary to the findings of [22], among the control variables shown on Table 4, lottery stock returns co-move positively with market capitalization and negatively with turnover.

Table 4 Determinants of return co-movement with lottery stock

This table reports the determinants of return co-movements with lottery stock estimated by following the regression during the first season 1997 to the first season 2011. The dependent variable is beta measures estimated by Eq. (4). Ins (mar) is the holdings of margin purchase traders (institutions). Index is lottery index. The remaining variables are the same as Table 1. Mean is the average of time-series coefficient estimates. Positive (negative) is the number of positive (negative) coefficients estimated quarterly.

$$Beta_{it} = \beta_0 LOTs_{it} + \beta_1 MValue_{it} + \beta_2 Turn_{it} + \beta_3 Lib_{it} + \beta_4 Eps_{it} + \beta_5 Saleg_{it} + \beta_6 \text{Re}\, turn_{it} + \beta_7 BM_{it} + \beta_8 Ins_{,t} + \beta_9 Ind_{it} + \beta_{10} Saleg_{it} + \beta_{11} LOTs_{it} * Ins_{it} + \beta_{12} LOTs_{it} * Ind_{it}$$

	index	size	turn	lib	eps	sale	ret	bv	ins	mar	index* ins	index* mar
mean	0.304	0.031	-0.089	0.031	-0.047	0.007	-0.041	0.038	0.135	-0.203	-0.125	0.242
positive	52	11	8	8	2	7	13	15	15	2	1	16
negative	0	1	22	1	10	4	21	8	1	13	16	1
p-value	0.000	0.016	0.001	0.000	0.002	0.507	0.112	0.071	0.000	0.000	0.000	0.000

3.5 Excess Return Of Lottery Stock

[9] proposed a three-factor model to capture the patterns in stock returns associated with risk premium (RP), size (SMB), and value versus growth (HML). Attempting to capture momentum returns (MON), [5] proposed a four-factor model for stock returns. Lottery stock captures the composite characteristics of small price, high volatility, and high skewness, not associated with SMB, HML, and MOM. This study examines whether empirical asset pricing models should capture the excess return of lottery stock, in addition to the four-factor model.

The literature ([6]; [8]) found that investors are willing to sacrifice their returns in the amount of a lottery premium, in exchange for low stock price. This work constructs a value-weighted portfolio return[5], the return difference between lottery stock and non-lottery stock portfolios in quarter t. The literature suggested that firm-size may serve as a proxy variable for various unobserved firm characteristics. The current study sorts firms in each quarter into five groups based on their market capitalization at the end of the prior quarter. We then investigate whether the model, which includes the excess return of lottery stock in addition to the Carhart's RP, SMB, HML, and MOM factors, can explain the cross-section

returns. All variables of the dependent variable and independent variables are standardized. The parameter estimates and test statistics are obtained from the time series of quarterly cross-sectional regression estimates as in [10]. The *p*-value for testing the significance of each coefficient is the *p*-value corresponding to the *t*-statistic calculated from the mean of the coefficient divided by its standard error.

Table 5 Excess return of lottery stock

This table reports the excess return of lottery stock during the first season 1997 to the first season 2011. The dependent variable is the value-weighted portfolio return sorted by market capitalization. RP is risk premium. SMB is value premium (small minus big). HML is value versus growth (high minus low). MOM is momentum returns. LOT-NONLOT is the return difference between portfolios of lottery and non-lottery stock. coe. is coefficient.

Market capitalization		RP	SMB	HML	MOM	LOT-NONLOT
small	coe.	0.027	0.071	-0.109	-0.034	0.251
	p-value	0.153	0.006	0	0.091	0
2	coe.	0.048	0.102	-0.101	-0.029	0.229
	p-value	0.014	0	0	0.169	0
3	coe.	0.053	0.107	-0.093	-0.024	0.209
	p-value	0.006	0	0	0.271	0
4	coe.	0.07	0.093	-0.095	-0.029	0.196
	p-value	0	0	0	0.196	0
high	coe.	0.064	0.058	-0.073	-0.038	0.16
	p-value	0.001	0.034	0	0.096	0

Table 5 presents the empirical results. In addition to RP, SMB, and HML, we find that the coefficients on lottery stock excess return are positive and significant at the 5% level among the five group stocks classified by market capitalizations. The absolute value of the coefficient of lottery stock excess return is largest among the five explaining factors, while the momentum factor loses influence on the five group portfolio returns. Lottery stock excess return monotonically decreases with market capitalization ranked by market capitalization of firms at the end of the prior quarter. The result suggests that the excess return of lottery stock captures some characteristics not related to RP, SMB, HML, and MON. To check the robustness of the effect of excess return of lottery stock on five sizes of portfolio returns, we also construct the equally weighted excess return of lottery stock and find a very similar pattern using this alternative measure.

3.6 Do Individual Investors Prefer Lottery Stock During An Economic Downturn?

Many researches assume that noise traders do not trade based on firm fundamentals, but in a herd-like manner. Individuals are generally considered as less informed than institutional investors, and margin purchases examined in this paper may be viewed as imperfectly informed speculative investors ([27]). During an economic downturn, investors might become frustrated and pessimistic and more willing to allocate more wealth to gamble-related stocks, which might provide a possible future fortune. If individual investors with strong gambling preferences discover the small possibility of a very large return, they might be

willing to allocate a larger portfolio of lottery stocks to eliminate poverty, even though lottery stocks are expected to underperform. This work investigates the time variation in the aggregate demand for lottery stocks by estimating the following time-series regression model:

$$EBSI_t = \alpha_1 M1b_t + \alpha_2 Mits_t + \alpha_3 Exp_t + \alpha_4 Mpi_t + \alpha_5 Uemp_t + \alpha_6 Pe_t + \alpha_7 Rp_t + \varepsilon_t \quad (6)$$

The dependent variable (*EBSI*) in the regression (6) is the excess buy-sell imbalance of lottery stocks traded by margin purchase traders in a given quarter *t*. EBSI is the buy-sell imbalance of a lottery stock portfolio minus the buy-sell imbalance of a portfolio that contains the other two categories of stocks (non-lottery and other stocks). EBSI measures individual investor preference toward lottery stocks related to the change toward the other two categories of stocks[6].

The independent variables in the regression include the following macroeconomic variables that may vary significantly over the business cycle: monetary aggregate (*M1b*), monitoring indicators (*Mits*), increased rate of export (*Exp*), industrial production index (*Mpi*), unemployment rate (*Uemp*), price-to-earnings ratio (*Pe*), and three month default risk premium[7]. All variables included in the regression are standardized. The available datasets reveal the holding of margin purchase traders used as proxy individuals, starting from the first season 1991; this study detect whether individual investors prefer lottery stock during an economic downturn from the first season 1991 to the first season 2011.

Table 6 Macroeconomic condition and lottery stock demand

This table reports the excess return of lottery stock during the first season 1991 to the first season 2011. The dependent variable is EBSI. EBSI is the buy-sell imbalance of a portfolio of lottery stocks minus the buy-sell imbalance of a portfolio that contains the two categories of stocks (non-lottery and other stocks). *M1b* is monetary aggregate. *Mits* is monitoring indicators. *Exp* is the increase rate of export. *Mpi* is the industrial production index. *Uemp* is the unemployment rate. *Pe* is the price-to-earnings ratio. *rp3m* is the 3-month default risk premium.

	coefficient	low	upper
M1B	0.145	-0.374	0.664
light	-0.087	-0.705	0.530
export	0.223	-0.268	0.714
MPI	0.018	-0.609	0.645
Uemp	0.278	0.030	0.526
pe	-0.261	-0.642	0.120
rp3m	0.030	-0.210	0.270

Table 6 reports the effect of economic downturn on individual investor trading behavior. Only the coefficient on unemployment rate is significant and positive at the 5% level, unlike the cases of other independent variables. Our results are consistent with the findings of [21]. The empirical results suggest that individuals tend to gamble during an economic downturn, implying that they are willing to risk earning a small chance of a huge gain at the expense of margin cost when they are in economic stress. Lottery-type stocks provide options to

individual investors with the opportunity to enrich their life in a declining market.

3.7 The Effect Of Lottery Stock On Market Returns And Volatility

A common belief is that investing in the stock market is more like gambling. If gambling was widely accepted and speculation in the stock market emerged, their systematic shifts could be associated with higher volatility and lower average return. That is, lottery stocks traded systemically by investors might influence stock market returns and volatility. The trading rationale of the speculator, who buys an asset at a lower price and sells at a higher price, may stabilize the market ([12]). On the contrary, positive feedback traders, who might be irrational speculators, will drive market volatility. The evidence of [16] shows that speculative trading activity through day trading increases intraday price volatility and exacerbates such volatility if the price rises on a particular day. [18] found no destabilizing effect of speculation in the post-World War II period or in the period prior to the 1929 crash, only a relation between speculation and volatility during 1934-1946. This study focuses on the effect of lottery stock, viewed as speculative objects, and non-lottery stock on market return and volatility.

Table 7 Lottery stock effect on market return and volatility

This table reports the lottery stock effect on market return and volatility during the third season 1986 to the first season 2011. Lotret and nonlotret are value-weighted lottery and non-lottery stock portfolio returns quarterly, respectively. mr is market return. turn is turnover.

$$r_t = \alpha_0 + \alpha_1 lotret_t + \alpha_2 nonlotret_t + \alpha_3 \log(h_t) + \varepsilon_t$$

$$\ln(h_t) = \beta_0 + \beta_1(\varepsilon_{t-1}/\sqrt{h_{t-1}}) + \beta_2\left[\left|\varepsilon_{t-1}/\sqrt{h_{t-1}}\right|\right] + \beta_3 \ln(h_{t-1}) + \beta_4 turn_t + \beta_5 rlotret_t + \beta_6 nonlotret$$

variable	coefficient	prob.	coefficient	prob.
	Mean equation			
α_0	0.385	0.150	0.461	0.078
lotret	0.450	0.000	0.447	0.000
nonlotret	0.565	0.000	0.573	0.000
$\log(h_t)$	0.496	0.004	0.518	0.001
	Variance equation			
β_0	-1.395	0.068	-1.408	0.048
$\varepsilon_{t-1}/\sqrt{h_{t-1}}$	0.090	0.504	0.141	0.243
$\lvert\varepsilon_{t-1}/\sqrt{h_{t-1}}\rvert$	-0.013	0.952	0.223	0.299
$\ln(h_{t-1})$	-0.593	0.000	-0.552	0.000
turn	0.015	0.056	0.014	0.066
lotret	0.015	0.189		
nonlotret	-0.044	0.018		
(mr>0)*lotret			0.018	0.097
(mr<0)*lotret			-0.006	0.771
(mr>0)*nonlotret			-0.063	0.008
(mr<0)*nonlotret			-0.018	0.361

To examine further the effect of lottery and non-lottery stock returns on market returns and volatility, the current study constructs quarterly value-weighted lottery and non-lottery stock portfolio returns. A large volume of literature has emerged on modeling and predicting volatility in financial markets because accurate volatility forecasts of financial assets are crucial inputs for derivatives pricing. This work uses the time-varying conditional variance of return as the proxy for market volatility and applies the EGARCH (1,1) process with GARCH-in-mean to estimate market volatility in this research.

For not decomposing market returns into positive and negative, Table 7 shows that value-weighted lottery and non-lottery returns are significantly and positively associated with market returns in the mean equation. No evidence supports that lottery stock returns could result in market volatility, or the destabilizing effect of lottery stock on market volatility. However, findings show that non-lottery stock returns have a negative effect on market volatility in the conditional variance equation.

As we further decompose market returns into up and down market returns, we find that lottery stock returns could weakly exacerbate market volatility during up market periods. Our evidence does not completely support irrational roles of lottery stock, while the negative effect of non-lottery stock returns on market volatility is found during the up market. Based on these empirical results, the evidence suggests a stabilizing effect of non-lottery stock on market volatility.

4 Conclusion

The stock market has often been compared to gambling. This study investigates lottery stock in the Taiwan Stock Market and finds market capitalizations and earnings negatively related to lottery stocks. The effect of gambling-induced stock on excess co-movements is significantly stronger on individuals than on institutions. Consistent with the findings of [21], our empirical results suggest that individuals tend to gamble during an economic downturn. Further decomposition of market returns into up and down market returns found that lottery stock returns could weakly exacerbate market volatility during up market periods. The negative effect of non-lottery stock returns on market volatility is only found during an up market, indicating that non-lottery stocks have a stabilizing effect on the market. Institutional investors increase the holdings of stocks and lottery stock favored only by individual investors, and we will like to see the weak effect of lottery stock on market volatility.

Reference

[1] Barber, B., Lee, Y.T., Liu, Y.J., Odean, T. *Just How Much do Individual Investors Lose by Trading*? Review of Financial Studies, 2009, 22, 609-632.
[2] Barberis, N. & Huang, M. *Stocks as Lotteries: The Implications of Probability Weighting for Security Prices*, American Economic Review, 2008, 98, 2066-2100.
[3] Boyer, B., Mitton, T. & Vorkink, K. *Expected Idiosyncratic Skewness*, Review of Financial Studies, 2010, 23, 169-202.
[4] Brunnermeier, M. K., Gollier, C. & Parker, J. A. *Optimal Beliefs, Asset Prices, and the Preference for Skewed Returns*, American Economic Review, 2007, 97, 159-165.
[5] Carhart, M. M. *On Persistence in Mutual Fund Performance*, Journal of Finance, 1997, 52, 57-82.

[6] Chen, M. H., Chen, S. J., Yen, M. F. & Shen, Y. C. *Lottery Premium in the Taiwan Stock Market*, Asia Pacific Management Review, 2008, 13(2), 545-556.
[7] Doran, J. S., Jiang, D. & Peterson, D. R. *Gambling Preference and the New Year Effect of Assets with Lottery Features*, Review of Finance, forthcoming.
[8] Downs, T.W. & Wen, Q. *Is there a Lottery Premium in the Stock Market*? Journal of Portfolio Management, 2001, 28, 112-119.
[9] Fama, E. F. & French, K. R. *Common Risk Factors in the Returns on Stocks and Bonds*. Journal of Financial Economics, 1993, 33, 3-56.
[10] Fama E. & MacBeth J. *Risk, Return, and Equilibrium: Empirical Tests*, 1973, Journal of Political Economy, 81(3), 607-636.
[11] Friedman, M. & Sayage, L. J. *The Utility Analysis of Choices Involving Risk*, Journal of Political Economy, 1948, 56(4), 279-304.
[12] Friedman, M. J. *Essays in positive economics*, 1953, Chicago University, Chicago press.
[13] Gao, X & Lin, T. C. *Do Individual Investors Trade Stocks as Gambling? Evidence from Repeated Natural Experiments*, 2011, working paper.
[14] Garrett, T. A. & Sobel, R. S. *Gamblers Favor Skewness, Not Risk: Further Evidence from United States' Lottery Games*, Economics Letters, 1999, 63(1), 85-90.
[15] Golec, J. & Tamarkin, M. *Bettors Love Skewness, not Risk, at the Horse Track*. Journal of Political Economy, 1998, 106, 205-225.
[16] Hsin, C. W., Guo, W. C., Tseng, S. S. & Luo, W. C. *The Impact of Speculative Trading on Stock Return Volatility: The Evidence from Taiwan*, Global Finance Journal, 2003, 14(3), 243-270
[17] Jadlow, J. W. & Mowen, J. C. *Comparing the Traits of Stock Market Investors and Gamblers*, The Journal of Behavioral Finance, 2010, 11, 67-81.
[18] Jones, J. D., Mulherin, J. H. & Titman, S. *Speculative Trading and Stock Market Volatility*, 1990, working paper.
[19] Kahneman, D. & Tversky, A. *Prospect Theory: An Anialysis of Decision Making under Risk*, Econometrica, 1979, 47(2), 263-291.
[20] Kumar, A. & Lee, C. M. C. *Retail Investor Sentiment and Return Comovements*, Journal of Finance, 2006, 61, 2451-2486.
[21] Kumar, A. *Who Gambles in the Stock Market*? The Journal of Finance, 2009, 44, 1889-1933.
[22] Kumar, A., Page, J. & Spalt, O. *Gambling and Comovements*, 2011a, Working Paper.
[23] Kumar, A., Page, J. & Spalt, O. *Religious Beliefs, Gambling Attitudes, and Financial Market Outcomes*, Journal of Financial Economics, 2011b,102(3),671-708..
[24] Markowitz, H. M. *The 'Two-Beta' Trap*, Journal of Portfolio Management, 1984, 11(1), 13-20.
[25] Odean, T. *Are Investors Reluctant to Realize their Losses*. Journal of Finance, 1998, 53(5), 1775-1798.
[26] Statman, M. *Lottery Players/Stock Traders*, Financial Analysts Journal, 2002, 58(1), 14-21.
[27] Stein, J. C. *Informational Externalities and Welfare-reducing Speculation*, Journal of political Economy, 1987, 95, 1123-1145.
[28] Thaler, R. N. & Johnson, E. J. *Gambling with the House Money and Trying to Break Even: The Effect of Prior Outcomes on Risky Choice*, Management Scinence, 1990, 36, 643-660.

[1] Gamblers and investors share five traits: material needs, competitiveness, superstition, financial conservatism (negative), and numeracy. They are different on three traits: present-time orientation, emotional instability, and impulsiveness.

[2] The Public Welfare lottery has been debuted since its inception in 2002 in Taiwan.

[3] Kumar (2008) chose the 30 and 50 percentile to identify lottery stocks.

[4] For each variable, its time-series average over the sample period is calculated and tested for significance, using a standard t-test in the following form : $a(\bar{\hat{r}}_i) = \bar{\hat{r}}_i / (\sigma(\hat{r}_i)/\sqrt{n})$, i=1,2,3…,n. where $\bar{\hat{r}}_i$ is the average of the quarter-by-quarter variable i, $\sigma(\hat{r}_i)$ is the standard quarterly deviation of variable i, and *n* is the total number of quarters.

[5] The results are persistent when we construct equal-weighted portfolio returns.

[6] According to the work of Kumar and Lee (2006), the buy-sell imbalance (BSI) of the lottery stock (non-lottery and other stocks) portfolio in given quarter *t* is defined as $BSI = 100(\sum_{i=1}^{n} BSI_i)/n$, where $BSI_i = (\sum_{j=1}^{d}(B_{ijt} - S_{ijt}))/((\sum_{j=1}^{d}(B_{ijt} + S_{ijt})))$. *d* is the number of days in a given quarter *t*, B_{ijt} is the buy volume (measured in dollars) for stock *i* on day *j* in quarter *t*, S_{ijt} is the sell volume for stock *i* on day *j* in quarter *t*, and *n* is the number of stocks in lottery stock (non-lottery and other stocks) portfolio.

[7] Three-month default risk premium denotes the interest rate of a commercial paper of 90 days minus the three-month term deposit interest rate.

EXAMINING THE IMPACT OF ISO 14000 CERTIFICATION ON SHAREHOLDER VALUE BY INDUSTRY SECTOR

Nabil Tamimi, tamimin1@scranton.edu, (570) 941-4288
Rose Sebastianelli, sebastianer1@scranton.edu, (570) 941-4287
Department of Operations and Information Management, Kania School of Management,
University of Scranton, 320 Madison Avenue, Scranton, Pennsylvania, 18510

ABSTRACT

We examine if improved environmental management impacts the economic bottom line by evaluating the stock prices of ISO 14000 certified companies by industry sector. We compare the annualized returns computed over various time periods (1-, 3- and 5- year) for ISO 14000 certified companies within a given sector to those of its corresponding S&P 500 sector index. In an effort to better mimic the way annualized returns are computed for the S&P 500 sector indexes, annualized returns for the ISO 14000 sector portfolios are also calculated using a weighting scheme based on market capitalization.

Keywords: ISO 14000, Stocks, Performance.

INTRODUCTION

Akin to the quality movement that began decades ago, environmental sustainability is poised to transform how companies do business. Parallels have been drawn between managing environmental sustainability and managing quality. First, international standards exist for the certification of both quality management and environmental management systems (ISO 9000 and ISO 14000). Second, the idea of designing quality into products and processes, rather than relying on end-of-process activities such as inspection, also applies to environmental sustainability. The best solutions are those that use eco-effective designs that focus on prevention rather than correction. And, finally, many argue that improved environmental performance, like improved quality, makes good business sense that affects the triple bottom line (social, environmental and economic).

In this paper, we examine if ISO 14000 certification, with its implied organizational commitment to improving environmental management, positively impacts a company's stock performance in the market (the economic bottom line). The ISO 14000 standards, developed in response to the United Nation (UN) goal of "sustainable development" discussed at the 1992 UN Conference on Environment and Development and revised subsequently in 2004, specify the requirements for an EMS (Environmental Management System). The standards provide the framework for a strategic approach to environmental management by not only requiring compliance with environmental legislation but a commitment to continuous improvement of the EMS. In order to evaluate how ISO 14000 certification may impact shareholder value , we create sector specific portfolios of ISO 14000 certified companies and compare their performance, in terms of annualized returns computed over 1-, 3- and 5-year periods, to those of their corresponding S&P

indexes. We also compare the performance of the overall S&P 500 index to the ISO 14000 portfolio including all sectors over the same holding periods.

RELEVANT LITERATURE

There has been anecdotal evidence demonstrating that companies seem to perform better financially when they manage environmental and social factors as well as economic factors. For example, Interface, Inc., a company that produces about 40% of the world's commercial carpet, saved a cumulative total of $222 million as a result of its environmental sustainability program [1]. In describing sustainability efforts undertaken at Ford, specifically the rebuilding of its Ford Rouge River Plant, the project manager notes "sustainability and quality serve the same purpose…companies who survive in today's global competitive marketplace are those that minimize waste including scrap, repair and rework; use the intellectual capacity of all their employees to drive continuous improvement and creative solutions; and support the communities in which they work and in which their customers dwell from a business, social and environmental perspective" [9, p.29]. By using recycled material in the production of parts, not only does Ford keep approximately 70,000 tons of waste from entering landfills, but also realizes cost savings (for example, saving $4.8 million in part costs by reusing crumb rubber).

While examples can be cited that show how improved environmental performance benefits companies, by reducing the costs of waste management, conserving energy and materials and/ or improving corporate image, do those benefits translate into improved profitability? Walley and Whitehead [10] argued that the cost of improving environmental performance actually reduces profitability and shareholder value. On the other hand, Porter and van der Linde [7] argued that improving environmental performance reduces waste and increases productivity therefore improving corporate performance. This debate is not unlike that which accompanied the quality (or TQM) movement when it began several decades ago.

Several empirical studies have suggested a positive relationship between improved quality and a company's profitability [2][8]. These studies investigated the linkage between attaining ISO 9000 certification and benefits to shareholders.

While examples of how environmental disasters have adversely affected stock prices do exist, such as the change in share prices of the Exxon Group after the Valdez accident [6] and more recently the drop (over 30%) in the value of BP shares following the Gulf disaster in April 2010, the linkage between improved environmental performance, overall profitability and shareholder value is less direct. Holt [4] conducted an exploratory study of UK companies that had achieved accreditation via BS 7750 environmental management standards. Her small-scale survey showed that 100% of the companies reported some degree of operational efficiency in waste minimization, 60% experienced a reduction in energy costs and 77% felt that employee morale was enhanced. However, none indicated a link between share price and BS 7750 accreditation. Feldman, Soyka and Ameer [3] provided a conceptual model linking environmental management systems (EMS) and environmental performance to market value for publicly traded corporations. Their model consisted of corporate EMS, environmental performance, environmental signaling, firm risk and firm value. After empirically testing their model using 300 of the largest publicly traded companies in the U.S., they found that improving both EMS and environmental

performance can increase a firm's stock price by as much as five percent. In a more recent study, Montabon, Melnyk, Sroufe and Calantone [5] conducted a large-scale survey of American managers' perceptions of the impact of ISO 14000 on corporate performance. Of the firms surveyed, 37% had attained ISO 14000 certification. Managers perceived ISO 14000 certification as having a positive impact on both the efficiency and effectiveness of the EMS and as greatly improving most dimensions of corporate performance including reduction of waste, company reputation and position in the marketplace.

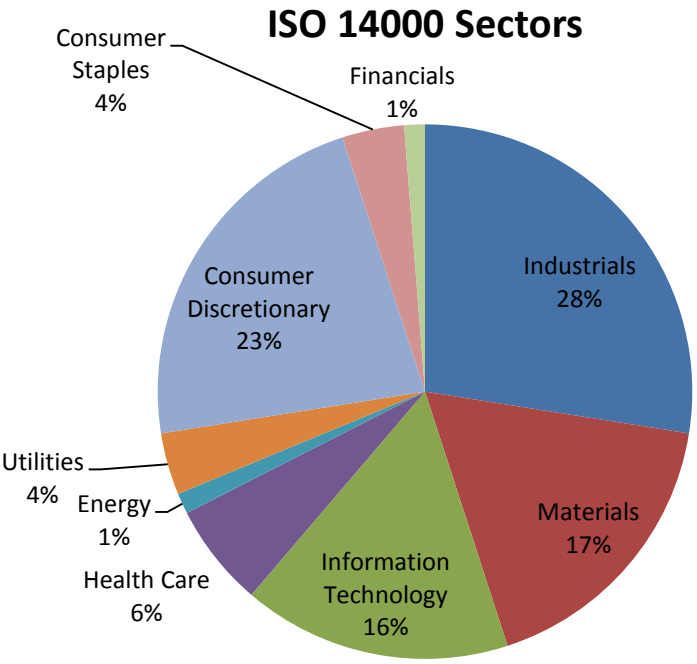
METHODOLOGY

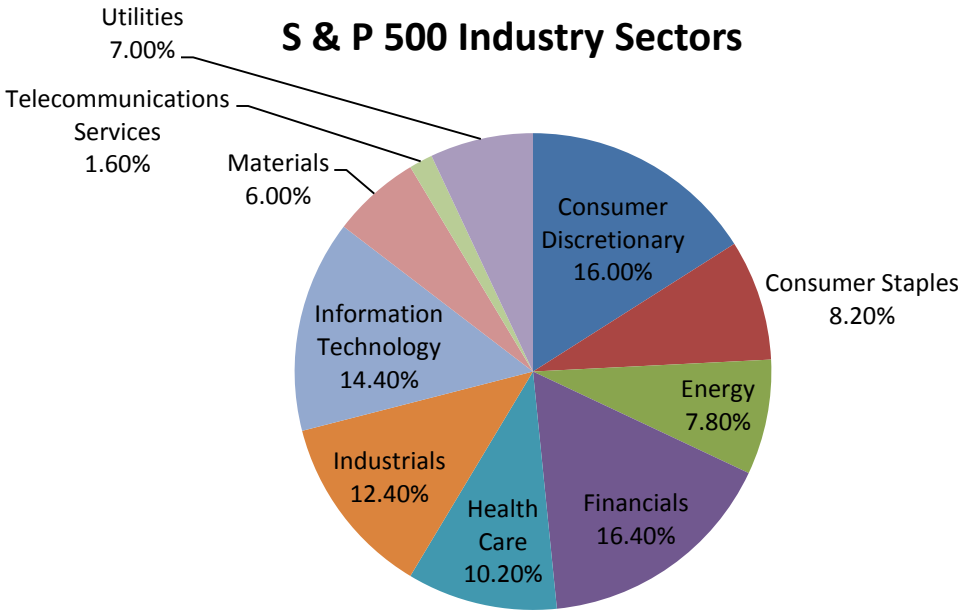
Our data set includes all publicly traded companies that achieved ISO 14000 certification during the period covered by the *ISO 14001 Worldwide Certified Company Directory—QSU Publishing Company, 2006*. To the best of our knowledge, no other comprehensive directory of this kind has been published since 2006. A total of 103 companies were listed as certified during this time period. Our sample consists of 79 companies; 24 were excluded because of acquisitions / mergers that occurred within the period covered by our study (through December 2010).

Historical monthly share prices (adjusted for dividends and splits) were obtained for each of the companies in our sample from *Yahoo! Finance* for the period from January 2006 through December 2010 (this period encompasses the recent financial crisis). Companies were grouped according to the Industry Classification Standard (GICS) taxonomy developed by Morgan Stanley Capital International (MSCI) and Standard & Poor's (S&P) for use by the global financial community. The GICS structure consists of 10 sectors: *Consumer Discretionary*, *Consumer Staples, Energy, Financials, Health Care, Industrials, Information Technology, Materials, Telecommunication Services* and *Utilities*. Since our ISO 14000 sample did not contain any companies that are part of the *Telecommunication Services* sector, no results are reported for this specific industry. The distribution of companies in each of these sectors for our ISO 14000 portfolio and the S&P 500 are shown in Figure 1.

We compute the annualized returns over various time periods (January 2010-December 2010, January 2008-December 2010, and January 2006-December 2010) for ISO 14000 certified companies within each industry sector. The annualized return follows the formula for the compound annual growth rate (CAGR) computed as: $(\text{Ending Value}/\text{Beginning Value})^{(1/n)} - 1$, where n = number of years. Moreover, the annualized returns for the ISO 14000 sector portfolios are also calculated for the 3 and 5 years' time periods using a weighting scheme based on market capitalization. This is done in an effort to mimic the way annualized returns are computed for the S&P 500 indexes where every stock is represented in proportion to its market capitalization.

FIGURE 1: COMPANY DISTRIBUTION BY INDUSTRY SECTOR





RESULTS

Table 1 compares the annualized returns computed over the three various holding periods (1-, 3- and 5- year) for the ISO 14000 portfolios within each industry sector to those of its corresponding S&P 500 sector index. Our results indicate higher 1-year annualized returns for

sector portfolios comprised of ISO 14000 companies compared to the S&P 500 sector index across all industries except *Health Care*. Annualized returns computed over a 3-year period are higher for ISO 14000 companies than for the S&P 500 in the *Energy*, *Industrials*, *Materials* and *Utilities* sectors. In the *Industrials* and *Utilities* sectors higher annualized returns for ISO 14000 companies are also realized over the 5-year period. The portfolio risk over its holding period is measured using the standard deviation. The standard deviation is a measure of the average spread of the sample of returns around the mean. If the returns are very spread out, the portfolio is labeled risky. Table 2 compares the annualized risk of the ISO 14000 portfolios to its corresponding S&P 500 across each industry sector for various holding periods. Our results show that the annualized risk of the ISO 14000 portfolios is much lower than the corresponding S&P 500 indexes across all industry sectors. The ISO 14000 annualized risk for the *Financials* and *Energy* sectors are not reported as each sector is comprised of a single company.

TABLE 1: ANNUALIZED RETURNS

	S&P 500 Annualized Returns			**ISO 14000 Annualized Returns**			**No. of S&P 500 Constituents**	**No. of ISO 14000 Constituents**
Sector	**1 Year**	**3 Years**	**5 Years**	**1 Year**	**3 Years**	**5 Years**		
Consumer Discretionary	27.66%	6.26%	4.32%	39.07%	4.19%	0.97%	80	18
Consumer Staples	14.11%	3.50%	7.68%	18.52%	-2.41%	4.72%	41	3
Energy	20.46%	-3.70%	8.31%	47.61%	-1.82%	4.15%	40	1
Financials	12.13%	16.25%	10.65%	38.56%	10.16%	-8.34%	81	1
Health Care	2.90%	-1.67%	1.84%	-0.41%	-8.62%	-6.11%	51	5
Industrials	26.73%	-2.72%	3.16%	38.17%	-1.64%	6.01%	58	22
Information Technology	10.19%	0.44%	5.02%	32.65%	-0.14%	-4.86%	76	12
Materials	22.20%	-0.45%	7.48%	36.51%	3.22%	5.76%	30	14
Utilities	5.46%	-5.71%	3.90%	10.84%	-3.95%	6.95%	34	3

TABLE 2: ANNUALIZED RISK

	S&P 500		**ISO 14000**	
Sector	**3 Years**	**5 Years**	**3 Years**	**5 Years**
Consumer Discretionary	27.34%	22.20%	18.14%	10.89%
Consumer Staples	14.53%	11.91%	5.06%	6.98%
Energy	25.56%	22.86%	n/a	n/a
Financials	38.18%	30.44%	n/a	n/a
Health Care	18.13%	15.27%	8.91%	10.58%
Industrials	29.24%	23.26%	10.84%	7.43%
Information Technology	26.33%	22.09%	11.08%	6.80%
Materials	30.40%	24.63%	17.08%	10.35%
Utilities	17.33%	15.27%	11.05%	6.05%

Annualized returns for the ISO 14000 sector portfolios are also calculated using a weighting scheme based on market capitalization to better mimic the way annualized returns are computed for the S&P 500 industry sectors. For each company in a given sector, we divide its market capitalization by the overall market capitalization of the entire sector under consideration. We use this proportion to determine how much to initially invest in a particular company and then use the annualized return of the company's stock price over its holding period to compute the ending value of the investment. We then compute the overall increase or decrease in the value of the portfolio for each industry sector in order to compute its corresponding annualized return. Table 3 illustrates the annualized returns for the weighted ISO 14000 portfolios compared to the S&P 500 industry sectors over 3-year and 5-year holding periods. Annualized returns computed over a 3-year period are higher for ISO 14000 companies than for the S&P 500 in the *Health Care*, *Industrials* and *Materials* sectors. In the *Consumer Staples*, *Health Care* and *Industrials* sectors higher annualized returns for ISO 14000 companies are also realized over the 5-year holding period.

TABLE 3: COMPARISON OF WEIGHTED ISO 14000 PORTFOLIOS WITH THE S&P 500 ANNUALIZED RETURNS

	S&P 500		**ISO 14000**	
Sector	**3 Years**	**5 Years**	**3 Years**	**5 Years**
Consumer Discretionary	6.26%	4.32%	-1.30%	-0.08%
Consumer Staples	3.50%	7.68%	2.04%	9.53%
Energy	-3.70%	8.31%	n/a	n/a
Financials	16.25%	10.65%	n/a	n/a
Health Care	-1.67%	1.84%	-1.16%	6.11%
Industrials	-2.72%	3.16%	-1.50%	4.66%
Information Technology	0.44%	5.02%	-2.13%	-4.03%
Materials	-0.45%	7.48%	3.38%	5.85%
Utilities	-5.71%	3.90%	-9.49%	3.54%

n/a: No data is reported as market capitalization information is not available.

To compare the performance of the broad S&P 500 index to the ISO 14000 portfolio representing all sectors, we compute the annualized returns over the same holding periods assuming equal dollar investment in each company comprising each sector. Table 4 illustrates the performance of the ISO 14000 portfolios compared to the S&P 500 index. The results demonstrate higher 1-year and 3-year annualized returns for the ISO 14000 portfolios relative to their corresponding S&P 500 index.

TABLE 4: COMPARISON OF ISO 14000 PORTFOLIOS TO THE S&P 500 INDEX

	Annualized Returns		
	1 Year	**3 Years**	**5 Years**
S&P 500 Index	15.06%	-2.80%	2.29%
ISO 14000 Portfolio	30.92%	-1.34%	0.58%

CONCLUSION AND IMPLICATIONS

Anecdotal evidence suggests that companies may perform better financially when they manage environmental and social factors as well as economic factors. To test this premise, we examine if ISO 14000 certification, with its implied organizational commitment to improving environmental management, positively impacts a company's stock performance in the market. In order to evaluate how ISO 14000 certification may impact shareholder value , we create sector specific portfolios of ISO 14000 certified companies and compare their performance, in terms of annualized returns computed over different holding periods, to those of their corresponding S&P 500 indexes. This approach provides us with a "micro-level" analysis of the annualized returns for each industry sector. We also create ISO 14000 portfolios using a weighting scheme based on market capitalization to better mimic the way annualized returns are computed for the S&P 500 industry sectors.

Our results regarding the non-weighted ISO 14000 portfolios indicate that investors are handsomely rewarded in the short term (i.e., January 2010-December 2010 holding period) as all industry sectors experienced significantly better annualized returns than their corresponding S&P 500 indexes; the only exception is the *Health Care* sector. Although the market overall has had quite a surge over the past year, our non-weighted ISO 14000 portfolios still significantly outperformed the S&P 500 indexes especially in the *Energy, Financials* and *Information Technology* sectors by 27.15%, 26.43% and 22.46%, respectively. One possible explanation for the underperformance of the *Health Care* sector may be attributed to investors' concern about major patent expirations in 2010. Pfizer, for instance, saw the patent protection of its blockbuster drug Lipitor and Aricept expire in 2010, representing $11.4 billion and $432 million, respectively, in 2009 revenues. For Merck, patents on Cozaar and Hyzaar drugs also expired in 2010 representing a combined $3.6 billion in revenues.

Annualized returns computed over a 3-year holding period (January 2008 – December 2010) are higher for the non-weighted ISO 14000 portfolios than for the S&P 500 indexes in the *Energy, Industrials, Materials* and *Utilities* sectors. However, it is interesting to note that during this holding period the annualized returns of our non-weighted ISO 14000 portfolios for *Consumer Staples, Energy, Health Care, Industrials, Information Technology and Utilities* were all negative. For the same time period, the S&P 500 indexes for *Energy, Health Care, Industrials, Materials* and *Utilities* also yielded negative annualized returns. Such results are not surprising as this period was part of the Great Crash of 2008-2009, where the broad S&P 500 index plunged 22.29%. Our weighted ISO 14000 portfolios outperformed the S&P 500 indexes in *Health Care*, *Industrials* and *Materials* for the same 3-year holding period, although two of the outperforming sectors (*Health care and Industrials*) still yielded negative annualized returns. And when compared to the broad S&P 500 index, the results demonstrate higher 1-year and 3-years annualized returns for the ISO 14000 portfolios relative to their corresponding S&P 500 index.

Although one implication of our findings is that building effective Environmental Management System (EMS) is ultimately rewarded by the market for certain industry sectors and holding periods, attributing shareholder benefits only to attaining ISO 14000 is difficult, and surely not the intent of this study.

REFERENCES

[1] Anderson, R. (2004) "Climbing Mount Sustainability," *Quality Progress,* Vol. 37, No. 2, 32-37.

[2] Docking, D. & Dowen, R. (1999) "Market Interpretation of ISO 9000 Registration," *The Journal of Financial Research*, Vol. 22, No. 2, 147-160.

[3] Feldman, S.J., Soyka, P.A., & Ameer, P.G. (1997) "Does Improving a Firm's Environmental Management System and Environmental Performance Result in a Higher Stock Price?" *Journal of Investing*, Vol. 6, No. 4, 87-97.

[4] Holt, D. (1998) "The Perceived Benefits of an Environmental Management Standard," *Business Process Management Journal*, Vol. 4, No. 3, 204-213.

[5] Montabon, F., Melnyk, S.A., Sroufe, R. & Calantone, R.J. (2000) "ISO 14000: Assessing Its Perceived Impact on Corporate Performance," *Journal of Supply Chain Management*, Vol. 36, No. 2, 4-16.

[6] Plesse, J. (1992) "Environmental Spending and Share Price Performance: The Petroleum Industry." *Business Strategy and the Environment*, Vol. 1, No.1.

[7] Porter, M. E. & van der Linde. C. (1995) "Green and Competitive: Ending the Stalemate." *Harvard Business Review,* 120-34.

[8] Rajan, M. & Tamimi, N. (2003) "Payoff to ISO 9000 Registration," *The Journal of Investing*, Vol. 12, No. 1, 71-77.

[9] Smith, L.R. (2004) "The Triple Top Line," *Quality Progress*, Vol. 37, No. 2, 23-31.

[10] Walley, N. & Whitehead, B. (1994) "It's Not Easy Being Green," *Harvard Business Review*, 46-52.

EXPLORING PAPER SUBMISSION PATTERNS OF ACADEMIC MULTIDISCIPLINARY CONFERENCES

Albert E. Avery, Department of Finance, Towson University, Towson, MD 21252
avery@towson.edu, 410-704-3185
Moon-Whoan Rhee, Department of Finance, Towson University, Towson, MD, 21252
mrhee@towson.edu, 410704-4075

ABSTRACT

Accurately forecasting conference attendance is very important in assuring the financial success of the conference. The most difficult part of determining attendance is determining the number of submissions received for presentations. Forecasting the total number of submissions for a conference has been shown by this research to be difficult and quite susceptible to large forecasting errors, but the error decreases as one progresses beyond the initial submission cutoff date. For the organizations under study, this research documents average forecasting error declining from approximately 27% at the initial submission cutoff date to zero 120 days later.

INTRODUCTION

Academic conferences have to exist and thrive for the academy to grow our knowledge, garner deeper insights into our respective areas of expertise and reach our potential. This research reports the progress of thought that follows new insights into data that has not been otherwise studied on an important subject: the financial success of a conference.

Academic conferences provide an arena to present new research ideas to an informed audience and to gain insight from colleagues outside the fairly strict accept/reject response from many journals. Some journal reviews returned to authors are insightful, while others are not. Conferences present the real-time opportunity to obtain feedback on thoughts and research in progress without the pain of rejection due to minor flaws or oversights.

The 18 conferences under study have benefitted from on-line paper submission, on-line blind review and on-line management, making them similar in character. The organizations under study conduct double or triple blind peer review as the basis for paper and special session acceptance.

Most conference participants present their research, as opposed to just attending presentation sessions. There are a number of others, particularly those who have been involved in organization leadership, who attend to remain abreast of current research in their field, network with colleagues and enjoy the social experience provided by receptions and other organized functions. Nevertheless, the number of submitted papers and PTWs heavily influences overall attendance. The total number of submissions to a conference is difficult to forecast, and forecasting this total is the focus of research reported here.

Serving as a Program Chair for academic conferences has many challenges, including the task of successfully encouraging faculty and practitioners to submit research papers, proposals to present in-progress research and proposals to offer panel, tutorial or workshop (PTW) sessions. The typical process of soliciting submissions begins with a call for papers (CFP) that specifies Tracks or areas of interest and a submission deadline. The CFP is typically emailed to past participants and also a marketing list.

All organizations in study purview track their conference participants, so it is easy to send a CFP. However, conference participants tend to "fall away" over time for a variety of reasons. In some cases, presentation of papers at conferences decreases in importance as a way of building a resume as the career of an author evolves. In other cases, the potential participant is attracted to alternative conference organizations. As such, sending a CFP to past participants only will result in conference attendance declining over time.

The conclusion that one draws from these observations from experience is that it is critically important for a Program Chair to be creative in constructing a list of potential submitters who have not submitted before. In addition, it is critically important to know when the rate of submissions is low and might lead to financial failure of the conference, so efforts to encourage submissions can be enhanced.

Other Research on These Data

Other research on conference data from these organizations has approached the probability of a specific person attending a conference as a Markov chain based on prior attendance [1]. Submission is predicted as a time-dependent coefficient Boolean dummy variables for submission in the prior year, plus coefficients times dummy variables for submission in earlier years. Coefficients for conferences further in the past than three years did not prove to be significant, so over the long run, conference attendance or paper submission rates can be viewed from other points of view. Specifically, the conferences themselves are viewed as independent events in this research.

THE STUDY

The Data and Objectives

In trying to optimize the number of submissions, setting due dates and extensions of due dates play a role in final submission counts, and hence acceptance counts and the ultimate financial success of the conference. Many conference expenses are fixed or semi-variable, so submission count shortfalls can have disastrous financial consequences. This research focuses on 2580 submissions to 18 academic conferences offered by three multidisciplinary academic organizations in the field of business between 2006 and 2011. The ultimate objectives are to determine whether it is possible to accurately forecast actual conference paper submission totals at the paper submission deadline, identify the extent of commonality in submission patterns and see the apparent effect of submission deadline extensions.

Table 1
Percentage of Total Submissions Received by the Cutoff Date

Organization and Meeting Year	At Cutoff
A-06	25.90%
A-07	26.22%
A-08	83.13%
A-09	19.53%
A-10	13.51%
A-11	38.16%
B-06	35.48%
B-07	19.07%
B-08	22.90%
B-09	20.65%
B-10	21.93%
B-11	22.66%
C-06	39.57%
C-07	44.52%
C-08	28.47%
C-09	13.87%
C-10	28.68%
C-11	42.00%
Overall	30.62%

To aptly compare conferences of different size, the data are recast as a percentage of total submissions received. It is apparent that the patterns of submission are somewhat similar, as shown by Table 1. The table reports the percentage of total submissions received at the end of the day on the initial submission cutoff date published in the CFP for organizations A, B and C.

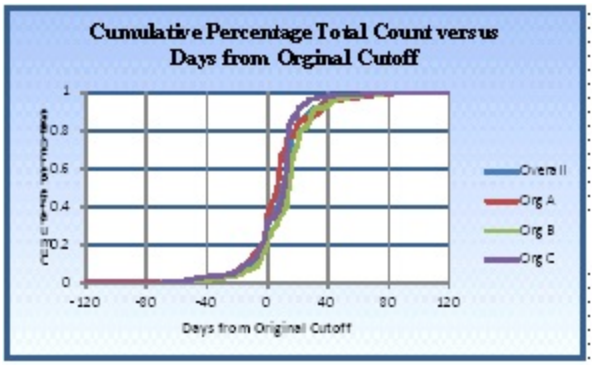
As implied earlier, it is common practice to extend the cutoff date one or more times for marketing purposes and it would appear that submitters expect the cutoff date to be extended. The overall pattern seems to be broken by Organization A in 2008 and 2011, Organization B in 2006 and Organization C in 2006, 2007 and 2011 by their notably higher percentage received figures.

Graph 1 reports averages for each of the three organizations, plus the overall averages. The overall averages are obtained by combining all 18 data sets. The X-axis reports the number of days between the initial or original submission deadline and the actual paper submission. Negative numbers represent days before the deadline. This graph reveals that very few papers are submitted earlier than one month prior to the initial deadline and, on the average, only 30.6% of submissions come in on time. As can be seen, the patterns are quite similar, even though

several conferences appear to be atypical, as shown below.

Graph 1

Overall and Organizational Cumulative Percentage versus Days from the Initial Cutoff Date



While it is normal to extend due dates by no more than two weeks, the graph shows that papers and proposals continue to be accepted for some time beyond, and submissions continue to come in more than 40 days after the initial submission cutoff date.

The diligent work of the respective program chair is credited to account for the submissions that arrive later than the extended cutoff date. It is not uncommon for a program chair to reach out to personally known colleagues, past and present officers and program committee members to solicit additional papers and especially PTWs. The time crunch makes blind reviewing of these submissions difficult, but still accomplishable.

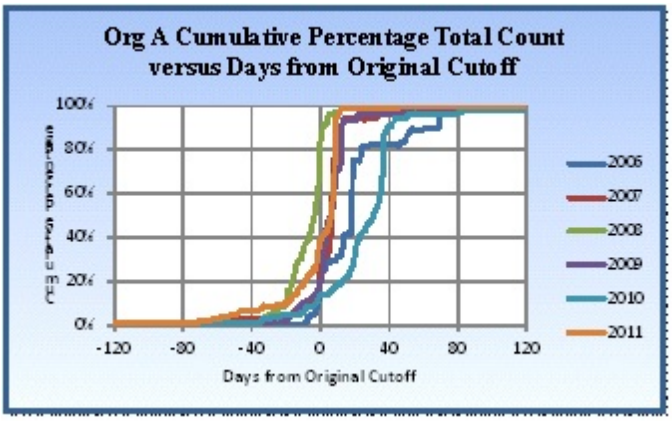
Graph 1 provides evidence that there may be a common pattern in submissions across all conferences that might provide for forecasting submission total counts. There are several ways this can be approached. There is a good deal of variation within organizations as shown in Graphs 2, 3 and 4

Regression Analysis

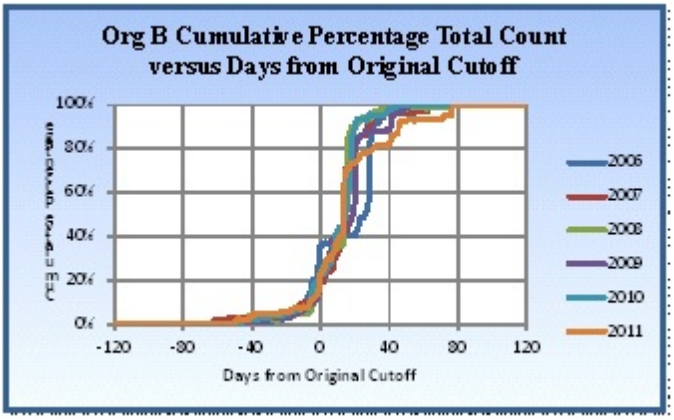
Common in financial research is the notion that returns for individual securities can be separated into normal returns and abnormal returns, with normal returns predicted using security beta and a market index. Here, it is possible that adjustments could be made for individual conference statistics using a conference beta and the overall submission percentages being used in place of a market index.

Table 2 reports the results of linear regressions on organizational submission percentages versus the combined overall percentages, while Graph 5 shows the relationships visually. There are two

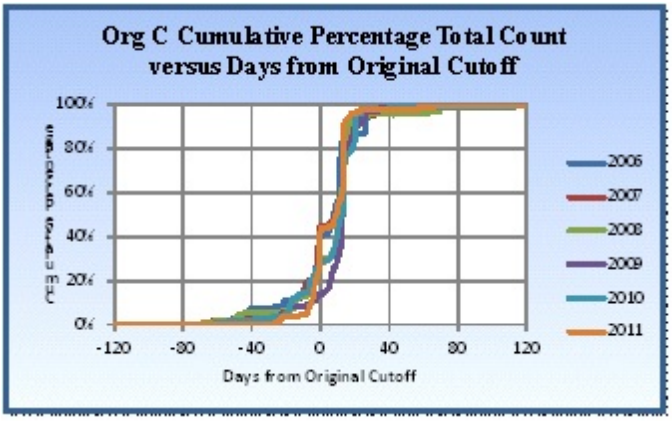
Graph 2
Organization A Conference Cumulative Percentages



Graph 3
Organization B Conference Cumulative Percentages



Graph 4
Organization C Conference Cumulative Percentages



tests of significance on beta, the slope. For the first test, the null hypothesis is that beta is zero, i.e. no relationship between the variables. A second test has a null hypothesis that the beta or

slope is 1.0. A slope of 1.0 implies that the changes in the organization percentages are equal to the changes in the overall percentages. If the alphas or intercepts were all equal to zero, a beta of 1.0 would assert that the organization percentages are equal to the overall percentages. Two of three alphas fail to pass this test. The critical t-test value for a 5% significance level is 1.9645 in all cases.

Table 2
Regression Results

Organization	Apha	Alpha t	Beta	Beta=0 t	Beta=1 t	Rsq
A	0.0118	4.26	0.9896	231.15	-2.43	99.55%
B	-0.0156	-4.89	0.9983	203.09	-0.35	99.74%
C	0.0026	1.19	1.0143	305.95	4.31	99.74%

Graph 5
Organizational Cumulative Percentages versus Overall Percentages

The calculated t-values for the beta = 0 hypothesis are incredibly far away from the critical value of 1.9645, so there is no statistical likelihood that the variables are unrelated. The calculated t-values for beta = 1 result in the null being rejected in two cases. Given the large sample size of 241 days, Organization A's beta of .9896 is statistically different from 1.0000. The same circumstance exists for Organization C, although this is not visually evident in Graph 4.

However, there is a difference between .9896 being statistically different from 1.0000 and the difference being of practical importance. it is apparent that, based on organizational averages, trying to adjust a final forecast for total submissions by using a beta for the organization will not change any practical results because the adjustment would amount to less than a 1% change in the forecast. Individual conference errors, reported below, are much larger.

This leads to tests on individual conferences. Using a "market index" that also includes the test conference as a component of the market index introduces a multicollinearity or variable problem between the regression variables that might reveal itself in regression residuals. The strength of the regression results suggest that this is not a serious concern.

While it would be desirable to be able to use the overall percentages as the "standard" in forecasting for individual conferences, it should not be undertaken without testing the importance of including or excluding a specific conference in the "market index." This leads to Tables 3 and 4. Highlighted values show where the null hypothesis is statistically (although not practically) significant.

In Table 3 and a critical t of 1.9645, the null hypothesis of Alpha = 0 is rejected with highlighted alpha values, which represent 13 of 18 regressions.. The null hypothesis of beta = 0 is soundly rejected in all cases. The null hypotheses that beta =1 is rejected in 13 of 18 highlighted cases. The smallest Coefficient of Determination is 90%, with nine values above 99%.

Table 3
Regression Results on Individual Conference Cumulative Submission Percentages against Overall Percentages with the Test Conference Excluded from the Overall Percentages

Conference	Alpha	Alpha=0 t	Beta	Beta=0 t	Beta=1 t	R-Square
1	**-0.0308**	-4.98	**0.9614**	101.18	-4.06	97.7%
2	0.0073	1.23	**1.0299**	113.15	3.29	98.2%
3	**0.0753**	5.56	0.9844	46.95	-0.75	90.2%
4	-0.0011	-0.20	**1.0288**	127.01	3.55	98.5%
5	**-0.0289**	-2.26	**0.9040**	46.32	-4.92	90.0%
6	**0.0506**	6.93	0.9991	88.64	-0.08	97.0%
7	-0.0142	-1.62	**0.9644**	71.66	-2.65	95.6%
8	**-0.0102**	-2.81	1.0062	180.96	1.12	99.3%
9	**-0.0198**	-4.67	**1.0268**	156.85	4.09	99.0%
10	**-0.0213**	-4.01	0.9992	122.11	-0.09	98.4%
11	**-0.0132**	-3.68	**1.0218**	184.39	3.93	99.3%
12	**-0.0082**	-2.23	**0.9604**	170.57	-7.03	99.2%
13	**0.0230**	6.43	0.9965	180.79	-0.63	99.3%
14	**0.0145**	4.92	**1.0126**	222.90	2.78	99.5%
15	**0.0103**	2.99	1.0005	187.96	0.09	99.3%
16	**-0.0187**	-3.99	**1.0314**	142.65	4.34	98.8%
17	0.0017	0.60	**1.0129**	226.33	2.88	99.5%
18	-0.0053	-1.29	**1.0373**	162.20	5.83	99.1%

In Table 4, the null hypothesis of alpha = 0 is rejected in 13 of 18 highlighted regressions. The null hypothesis of beta = 0 is soundly rejected in all cases. The null hypotheses that beta =1 is rejected in 12 of 18 highlighted cases. The smallest Coefficient of Determination is 91%, with 10 values above 99%.

The results obtained by using all conferences in the average is strikingly similar to results obtained by excluding the conference under test from the overall average, so further efforts will use the overall average results in seeking forecasting solutions.

Naive Forecasting

Since the patterns of submission appear to be quite similar and the focal point has been the initial submission cutoff date, a naive forecast for total submissions could be obtained by recognizing that in the overall averages, 30.62% of the submissions were received by that date. One can then extrapolate from the actual submission count at the cutoff date.

Table 4
Regression Results on Individual Conference Cumulative Submission Percentages against Overall Percentages with the Test Conference Included from the Overall Percentages

Conference	Alpha	Alpha=0 t	Beta	Beta=0 t	Beta=1 t	R-Square
1	**-0.0297**	-5.06	**0.9646**	107.04	-3.92	98.0%
2	0.0064	1.15	**1.0292**	120.13	3.41	98.4%
3	**0.0686**	5.35	0.9907	49.97	-0.47	91.3%
4	-0.0014	-0.28	**1.0280**	134.81	3.67	98.7%
5	**-0.0297**	-2.44	**0.9137**	49.05	-4.63	91.0%
6	**0.0470**	6.83	1.0008	94.00	0.07	97.4%
7	-0.0145	-1.74	**0.9686**	75.91	-2.46	96.0%
8	**-0.0098**	-2.87	1.0063	191.74	1.19	99.4%
9	**-0.0189**	-4.74	**1.0258**	166.42	4.18	99.1%
10	**-0.0205**	-4.09	1.0001	129.40	0.02	98.6%
11	**-0.0126**	-3.73	**1.0209**	195.55	4.01	99.4%
12	**-0.0079**	-2.28	**0.9629**	180.29	-6.94	99.3%
13	**0.0215**	6.38	0.9971	191.47	-0.56	99.4%
14	**0.0135**	4.88	**1.0122**	236.24	2.84	99.6%
15	**0.0096**	2.94	1.0008	199.10	0.16	99.4%
16	**-0.0179**	-4.06	**1.0303**	151.40	4.45	99.0%
17	0.0015	0.56	**1.0124**	239.88	2.94	99.6%
18	-0.0053	-1.35	**1.0357**	172.19	5.93	99.2%

Graph 6 reports the forecasting error for the 18 individual conferences and it is apparent that this forecasting method is quite unreliable. In two cases, the forecast was below the actual count by more than 120%. In one case, the forecast was above the actual count by 63%.

As shown in Table 5, even forecasting within organizations produces significant errors.

Improving Reliability in Forecasts

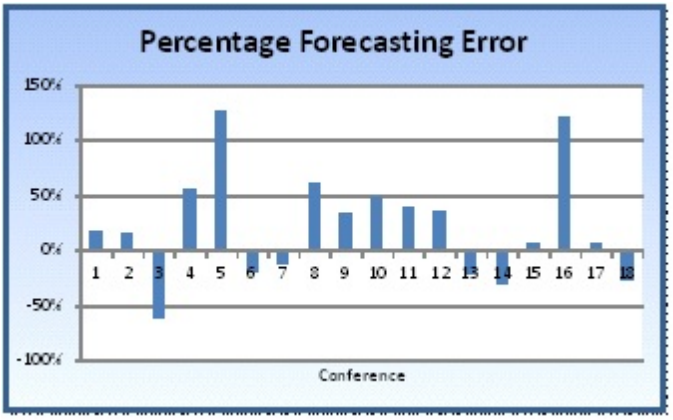
As one moves beyond the initial submission cutoff date and a higher proportion of total submissions have come in, it is logical that forecasts of the total submission count would improve. All submissions were received within 120 days of the initial cutoff date, so there would be little doubt about an accurate final count estimate on day 118 or 119. Uncertainty which cannot be mitigated increases as one moves backward toward the initial cutoff date.

Table 5
Within Organization Forecasting
Based on Count at the Cutoff Date

	Overall	Org A	Org B	Org C
Count	790	336	197	257
Forecast		1097	643	839
Actual	2580	946	846	788
Error		-151	203	-51
% error		-13.79%	31.50%	-6.11%

As noted earlier, the submission cutoff date is commonly extended, so another point of interest could be the final submission cutoff date. Tests to see whether naive forecasting from the final submission cutoff would improve the results did not bear fruit. The suspected reason is that specifically requested submissions would come in and be accepted by the program chair well after the final submission cutoff date. This is particularly true of PTWs.

Graph 6
Naive Forecasting Errors on the Submission Cutoff Date



The Remaining Forecasting Option

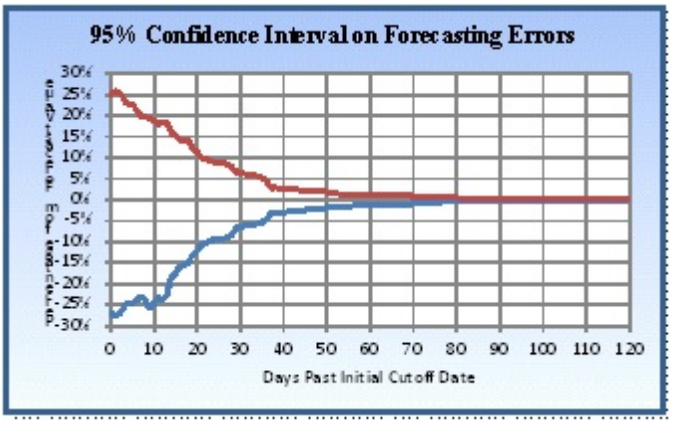
The remaining option for forecasting is to produce a new forecast each day based on submissions received to that point in time, extrapolating from the overall average on the proportion of papers yet to be received. This would provide a forecast on Day 120 with zero error. Logically, the forecasting error should be highest at the initial cutoff date and should diminish according to some pattern as one approaches Day 120. Using 20-20 hindsight, this pattern can be discerned.

The forecasting errors for each day for each of the 18 conferences form 200 sampling distributions of error terms, one for each day, with 18 observations each. Student t-tests were performed on the mean error for each day. Based on a 5% significance level in a two-tailed t-test, the null hypothesis that the mean error = 0 was not rejected in any case. As can be seen in Graph 7, the standard deviation of the errors declined dramatically during the first 37 days and continues to decline beyond that point.

Approaching forecasting accuracy from the point of view of a confidence interval, the calculated results for a 95% confidence interval are reported in Graph 8, with actual data reported in Table 6. Initially, the upper bound is 25.1% and the lower bound is -26.9%, so if the forecast total count on Day 0 was 100, the confidence interval would run from 73.1 to 125.1.

Graph 7
Standard Deviation of Forecasting Errors

Graph 8
95% Confidence Interval as a Percentage of the Forecast Submission Count



By Day 21, the interval is approximately the forecast value plus or minus 10%. By Day 36, the interval is just a bit better than plus or minus 5%. Table 6 shows the average of the absolute values of the upper and lower confidence interval limits essentially decreasing monotonically from 27% to zero as one approaches 120 days beyond the initial submission cutoff date.

CONCLUSION

Accurately forecasting conference attendance is very important in assuring the financial success of the conference. The most difficult part of determining attendance is determining the number of submissions received for presentations. Forecasting the total number of submissions for a

conference has been shown by this research to be difficult and also quite susceptible to large forecasting errors.

Table 6
Confidence Interval Limits and Absolute Averages

95% confidence interval			
Day	Low Limit	High Limit	Avg Abs
0	-26.86%	25.08%	26.0%
1	-27.68%	25.80%	26.7%
2	-27.09%	25.40%	26.2%
3	-25.51%	23.82%	24.7%
4	-24.49%	22.76%	23.6%
5	-24.68%	22.75%	23.7%
6	-23.80%	21.28%	22.5%
7	-23.30%	19.77%	21.5%
8	-24.12%	19.57%	21.8%
9	-25.81%	19.43%	22.6%
10	-24.56%	18.68%	21.6%
11	-23.48%	18.05%	20.8%
12	-23.56%	18.42%	21.0%
13	-22.62%	18.21%	20.4%
14	-18.53%	16.08%	17.3%
15	-17.56%	15.26%	16.4%
16	-16.09%	14.24%	15.2%
17	-15.61%	13.98%	14.8%
18	-14.86%	13.77%	14.3%
19	-12.98%	12.12%	12.6%
20	-12.14%	11.33%	11.7%
21	-10.77%	9.96%	10.4%
22	-10.42%	9.71%	10.1%
23	-9.79%	9.13%	9.5%
24	-9.53%	8.93%	9.2%
25	-9.47%	8.89%	9.2%
26	-9.27%	8.71%	9.0%
27	-9.01%	8.32%	8.7%
28	-8.42%	7.89%	8.2%
29	-6.79%	6.68%	6.7%
30	-6.62%	6.54%	6.6%
31	-6.23%	6.16%	6.2%
32	-6.08%	6.00%	6.0%
33	-5.94%	5.86%	5.9%
34	-5.74%	5.58%	5.7%
35	-5.37%	5.22%	5.3%
36	-4.61%	4.45%	4.5%
37	-3.16%	2.92%	3.0%
38	-3.11%	2.83%	3.0%
39	-3.02%	2.77%	2.9%
40	-3.00%	2.78%	2.9%
41	-2.84%	2.62%	2.7%
42	-2.58%	2.35%	2.5%
43	-2.55%	2.34%	2.4%
44	-2.47%	2.27%	2.4%
45	-2.44%	2.26%	2.4%
46	-2.21%	2.07%	2.1%
47	-2.06%	1.92%	2.0%
48	-2.08%	1.92%	2.0%
49	-2.11%	1.94%	2.0%
50	-1.93%	1.78%	1.9%
51	-1.76%	1.61%	1.7%
52	-1.63%	1.47%	1.5%
53	-1.53%	1.35%	1.4%
54	-1.53%	1.35%	1.4%
55	-1.50%	1.34%	1.4%
56	-1.45%	1.31%	1.4%
57	-1.45%	1.31%	1.4%
58	-1.43%	1.34%	1.4%
59	-1.43%	1.35%	1.4%
60	-1.43%	1.35%	1.4%
61	-1.37%	1.28%	1.3%
62	-1.37%	1.29%	1.3%
63	-1.42%	1.24%	1.3%
64	-1.42%	1.23%	1.3%
65	-1.42%	1.23%	1.3%
66	-1.42%	1.23%	1.3%
67	-1.41%	1.26%	1.3%
68	-1.41%	1.26%	1.3%
69	-1.41%	1.26%	1.3%
70	-0.84%	0.73%	0.8%
71	-0.83%	0.75%	0.8%
72	-0.68%	0.62%	0.6%
73	-0.68%	0.62%	0.6%
74	-0.68%	0.62%	0.6%
75	-0.68%	0.62%	0.6%
76	-0.68%	0.62%	0.6%
77	-0.43%	0.43%	0.4%

78	-0.43%	0.43%	0.4%
79	-0.43%	0.43%	0.4%
80	-0.43%	0.43%	0.4%
81	-0.36%	0.35%	0.4%
82	-0.36%	0.35%	0.4%
83	-0.36%	0.35%	0.4%
84	-0.31%	0.29%	0.3%
85	-0.27%	0.26%	0.3%
86	-0.27%	0.26%	0.3%
87	-0.27%	0.26%	0.3%
88	-0.27%	0.26%	0.3%
89	-0.27%	0.26%	0.3%
90	-0.26%	0.25%	0.3%
91	-0.26%	0.25%	0.3%
92	-0.26%	0.25%	0.3%
93	-0.26%	0.25%	0.3%
94	-0.26%	0.25%	0.3%
95	-0.26%	0.25%	0.3%
96	-0.26%	0.25%	0.3%
97	-0.26%	0.25%	0.3%
98	-0.26%	0.25%	0.3%
99	-0.26%	0.25%	0.3%
100	-0.26%	0.25%	0.3%
101	-0.26%	0.25%	0.3%
102	-0.26%	0.25%	0.3%
103	-0.26%	0.25%	0.3%
104	-0.26%	0.25%	0.3%
105	-0.26%	0.25%	0.3%
106	-0.26%	0.25%	0.3%
107	-0.25%	0.25%	0.2%
108	-0.25%	0.25%	0.2%
109	-0.25%	0.25%	0.2%
110	-0.25%	0.25%	0.2%
111	-0.25%	0.25%	0.2%
112	-0.25%	0.25%	0.2%
113	-0.25%	0.25%	0.2%
114	-0.23%	0.24%	0.2%
115	-0.23%	0.24%	0.2%
116	-0.23%	0.24%	0.2%
117	-0.23%	0.24%	0.2%
118	-0.23%	0.24%	0.2%
119	-0.23%	0.24%	0.2%
120	0.00%	0.00%	0.0%

The degree of uncertainty in a naive forecast model decreases as one progress past the initial submission deadline date. Each day presents a new total for submissions actually received, which provides for a new forecast of total submissions. Each day past the initial submission cutoff date improves the accuracy of the forecast. Specifically, average forecasting error decreases from approximately 27% to zero across the 120 day horizon, thus documenting how confidence in the forecast increases by day in the resulting forecast. This will help program chairs to better know how their respective conferences are progressing.

REFERENCES

1. Avery, A., *Patterns of Participation in Northeast Decision Sciences Institute Conferences.* Northeast Decision Sciences Institute Proceedings, April 1996, 368-389.

CURRENT CHALLENGES FOR EMPLOYEES IN PRIVATE EMPLOYER SPONSORED RETIREMENT PLANS

James Bishop, Bryant University, Suite C217 Unistructure, 1150 Douglas Pike, Smithfield RI 02917, jabishop@bryant.edu, 401-232-6356

ABSTRACT

The focus of this paper is to examine recent retirement benefit issues arising from plans that derive benefits from account balances. Many private employers have changed their retirement plan from a traditional defined benefit (DB) plan to a "hybrid" plan. A hybrid plan earns benefit credit in the form of a lump sum rather than a monthly life annuity. At the same time, interest rates used in converting a balance to an annuity (or vice versa) has changed from fixed treasury rates to segment rates based on corporate bond yield curves. As a result, hybrid plan retirement annuities are subject to large variability. This is an undesirable result for the employee. A simulation model is used within to model the variability of annuities payable for two types of defined benefit plans.

Keywords: retirement benefits, cash balance, stable value, pension equity, life annuity

INTRODUCTION\BACKGROUND

In recent years, many private employers have transitioned from offering traditional defined benefit (DB) retirement plans to "hybrid" plans. A hybrid plan is legally considered a defined benefit plan, but has the balance accumulation feature of a defined contribution plan (such as a 401k plan). Hybrid plans have a variety of names such as: "cash balance", "pension equity", or "stable value" [2]. These hybrid plans are completely employer funded, but unlike defined contribution plans that typically invest the contributions into the stock market, hybrid plan contributions earn stable growth defined by the plan [8] [9].

The benefit in a traditional defined benefit plan is based on a formula, computed upon retirement (or termination of employment) and based upon length of service and/or pay and is payable as a life annuity [5]. In contrast, the main feature of a hybrid plan is that the money available to the individual at retirement is a lump sum. The lump sum benefit is then converted to an actuarially equivalent life annuity (taking interest and mortality into account). This conversion factor is particularly sensitive to the interest rate in effect at the time of conversion, and these rates typically change annually.

We will focus on two types of hybrid pension plans within. First, a cash balance plan, which is an account balance that is supported by contributions from the employer as a percentage of the employee's salary [7]. The contribution percentage of pay may increase as an incentive to older or longer service employees as well. Contributions are added to the account periodically, typically either monthly or quarterly. This is a notional account that earns either a fixed interest rate or one that varies based on a relatively stable economic rate (such as the 30-year U.S. treasury rate).

Upon retirement from a company that sponsors a cash balance plan, the employee may or **may not** be eligible to take his or her cash balance as a lump sum. Pension plans must offer the employee a single life annuity if the employee is not married at the time of payment. Or, if married, the employee is offered an annuity benefit with at least 50% of the benefit going to the spouse in the event that the employee dies prior to the spouse.

For many cash balance pension plans, the cash balance (or possibly an amount more than the cash balance amount) may be taken as a lump sum at the employee's termination of employment. This money may then be rolled over to another retirement vehicle (IRA), particularly appealing for employees leaving a company prior to typical retirement age. The decision to take a lump sum benefit or life annuity benefit is an important one for employees who are preparing for retirement. The payment form chosen cannot be changed once elected.

The second type of hybrid plan we consider is a pension equity plan (stable value plan is another name for this plan type) For this plan type, the contribution credit is usually expressed as a percentage of pay, service, or age and the interest credit is usually defined as an indexed or fixed rate. Pension equity plans express the employee's benefit as a lump sum amount **payable at age 65**, not immediately. Contributions are indexed based on increases in participant compensation, rather than a specified interest rate (Watson Wyatt, 2003).

Table 1 summarizes the key distinctions between a traditional defined benefit plan, cash balance plan, and pension equity plan.

Table 1 – Summary benefit features for pension plan types

	Final average pay plan	**Cash balance plan**	**Pension equity plan**
Benefit formula result	Life annuity	Lump sum balance	Lump sum balance
Age of assumed payment	Normal retirement date – age 65	Immediate	Normal retirement date – age 65
Interest growth	None	Fixed rate or conservative economic rate	Wage indexed
Conversion to alternate form of payment	Lump sum – PPA interest and mortality	Life annuity – PPA interest and mortality	Life annuity – PPA interest and mortality

As part of the Pension Protection Act (PPA) of 2006, lump sum conversion interest rates and mortality were regulated. The required rates used are segment rates over three time periods based on corporate bond yield curves. The first rate is used for all discounting during the first five years, the second rate is applicable between years five and twenty, and the third rate is applicable for periods beyond twenty years.

As an example of the impact of benefit conversions, suppose an employee is age 64 and has a cash balance benefit of $500,000 payable immediately. The equivalent annual life annuity computed using an interest rate of 5% is $40,393 (the single life annuity factor is computed using the Pension Protection Act 2012 unisex mortality table). Instead, assume that this same employee works until age 65 and has a cash balance benefit of $515,000 payable immediately (assuming a 3% increase in benefit). If the interest rate at age 65 has changed to 6%, then the resulting annual life annuity is $39,042. Working an extra year has **reduced** this retiree's life annuity income by 3.3%. This situation is not possible under a traditional defined benefit plan.

Employers typically send benefit statements to the employees (required by law every three years), showing current and projected future benefits. These projections are simplistic and assume constant yield, projected wage increases, and interest assumptions. These projections are likely to be quite different from the benefit actually payable at retirement. This poses a serious challenge when considering retirement.

SIMULATION MODEL

Rather than using the typical constant interest rate assumptions that are commonly used to project and convert benefits for a future retirement date, simulation techniques allow us to project benefits in a more realistic manner. In addition, we can compare and contrast variation in the single life annuity that is likely to be paid at retirement for three common types of defined benefit plans.

In practice, each of the three defined benefit plan types discussed here have numerous specific formula designs depending on the goals of the employer in establishing the plan. For the purpose of comparing the basic features of each of the three plan types, a number of plan features and employee characteristics have been held constant across the three plan types. First, we consider full careers (40 years of service) and a constant wage increase that is the same for all three plans. Also, the choice of the specific benefit multipliers have been set to make each of the three plans give approximately the same annual annuity benefit.

The traditional final average pay plan has an arbitrarily set plan formula at with a 1.5% multiplier. This value is considered one possible benchmark for the industry because when multiplied by 40 years of service, gives an annual benefit to a retiree that is approximately 60% of their final average pay. Combining with social security, this benefit is considered a reasonable retirement income.

The initial goal of the simulation presented within is to set the multipliers for the cash balance plan formula and pension equity plan formula to be such that the annual annuity benefit is approximately the same of each other and the final average pay benefit as well. This serves two purposes, the first of which is to see how generous the benefits are for the three plans relative to each other. Also, our main goal of the simulation is to see how much variability is in the annual annuity result. Therefore, we hold the median of the three plan annuity values approximately constant across the three plans.

General assumptions for each model:

- wage increase 3.8% (last 20 year geometric average of the U.S. national average wage index)
- starting pay $50,000 at age 25
- retire at age 65

The following models are used to compare the three defined benefit plans under consideration:

Final Average Pay Plan:

- benefit formula = 1.5% x years of service x FAP5; where FAP5 is the final five years of annual pay prior to retirement.

  benefit payable as an annual annuity = constant value $119,406 per year

No simulation necessary for the final average pay plan as the annuity is fixed once the wages are known. Since all three benefit plans are based on the wage increases, these values remain constant across all benefit plans for the purpose of a fair comparison.

Cash Balance Plan:

- benefit formula =
- cb% = 8.25% = cash balance pay multiplier
- contributions are made mid-year

  simulation of i_k = annual interest rates utilize discrete uniform random selection based on 30-year U.S. treasury rates from 1988 – 2010.
  simulation of v is used for the conversion to an annuity as part of the factor **,** based on discrete uniform random selection of annual corporate moody (Aaa) bond yield from 1976 – 2010.

  benefit payable as an annual annuity = benefit formula ÷

  Where = P(person age 65 survives for k more years).

The cash balance simulation has two parts. The first part is to project the forty year future values for "typical" fluctuation in the interest credit rate applied to the cash balance account. Secondly, once the final balance is known at retirement, simulation is performed to convert the final cash balance account to an annual single life annuity.

Pension Equity Plan:

- benefit formula =
- PEP% = 13.25% = pension equity plan pay multiplier

- contributions are made at the end of each year
- ss_k = annual wage increase 3.8% used to annually increase balance

The benefit formula simplifies to \$6,000 x 40 x 1.038^{39} = \$1,134,861.

simulation - conversion to annuity using factor discrete uniform random selection based on annual corporate bond yield from 1976 – 2010.

benefit payable as an annual annuity = benefit formula ÷

The pension equity plan simulation only produces one value per run. Once the final balance is known at retirement, simulation is performed to convert the final pension equity balance to an annual single life annuity.

An initial series of sets of 1,000 simulations is performed using the Excel random number generator and discrete uniform selection of historical 30-year treasury rates and corporate bond yield rates. For each simulation, 40 years of projected 30-year treasury rates is produced and the final age 65 cash balance lump sum benefit determined. Then, one corporate bond yield rate is simulated to create an actuarial factor that converts the lump sum benefit to an annual life annuity benefit.

The initial simulations are run using various cash balance formula multipliers (6% – 10% by 0.25% increments) and the median resulting annual annuity closest to the final average pay plan annuity is chosen. For the same simulation numbers, the pension equity formula multiplier (10% - 15% by 0.25% increments) is similarly chosen.

Finally, a set of 1,000 simulations is performed in order to determine the standard deviation of resulting annuities for both the cash balance plan and the pension equity plan.

RESULTS

The initial sets of simulations produced the Key Formula Benefit Levels for the cash balance and pension equity plans as displayed in Table 2. The cash balance plan value of 8.25% produced a median annual single life annuity benefit closest to the final average pay plan value of \$119,406. Similarly, the pension equity plan value of 13.25% produced a median annual single life annuity benefit closest to the final average pay plan value of \$119,406.

A final set of 1,000 simulations produced benefit results for the cash balance plan and pension equity plan. A summary of results for life annuity benefits payable at age 65 is displayed in Table 2.

Table 2 – Summary single life annuity data for pension plan types

	Final average pay plan	Cash balance plan	Pension equity plan
Key Formula Benefit Levels	1.5%	8.25%	13.25%
Median Single Life Annuity	$ 119,406	$ 117,191.34	$ 119,024.84
Standard Deviation of the Median	$ 0	$ 23,157.38	$ 22,061.04
Standard Deviation as Percentage of the Median	N/A	19.8%	18.5%

Assuming a stable projected wage increase the final average pay plan formula chosen here results in a constant annual single life annuity benefit of $119,406. In order to approximately match the final average pay annuity, the cash balance plan multiplier of 8.25% is required. This multiplier is quite high as compared to existing cash balance plans in the industry. Typical multipliers are in the range of 5.0%. Likewise, the pension equity plan multiplier is determined to be 13.25%. This value is moderately high as compared to commensurate plans in the industry. Typical multipliers are in the range of 10.0%.

The standard deviation (SD) associated with the cash balance plan is close to 20% relative to the median projected single life annuity benefit. This may be a cause of concern for an employee planning for retirement under this plan type. The pension equity plan has an 18.5% standard deviation as a percentage of the median annuity benefit. This is a surprising result because unlike the cash balance plan, the only source of variation is the life annuity factor used to convert the lump sum to an annuity.

Both the cash balance and pension equity plans used the same simulation numbers to convert the final age 65 lump sum amounts to life annuity values. Therefore this source of variation is the same for both plans. The small difference between standard deviations for the two plans is created by the simulated projection of the cash balance plan interest rates. This projection created very little additional variation in results.

CONCLUSION/DISCUSSION

Retirement planning is an area of concern in recent years. Employers are eliminating or converting traditional defined benefit plans into benefits derived from account balances. A stable life annuity promise offered through the employer has almost become a thing of the past.

Combined with this trend, the economic downturn in 2007 has added to retirement instability because of the sharp stock market decline. This has created a mistrust of account balance derived benefits. This paper confirms the mistrust of benefit stability, but in the case of cash balance and pension equity plans the culprit for benefit instability is interest rates used to convert the balances to life annuities.

One major purpose of recent account balance pension designs has been to stabilize account balance benefits so that they are not subject to the wide market variability that exists in defined contribution (401k) plans. Ironically, due to the interest rates used in the conversion from an account balance to a life annuity, the resulting annuity is unstable at the point that an employee is actually choosing to take their benefit.

As a final perspective for employees who are concerned with receiving an adequate life annuity benefit, we consider the differences between traditional defined benefit, new defined benefit (cash balance and pension equity), and 401k plan offerings. If we keep wage increases constant across plan types then the traditional defined benefit plan produces a fixed annuity – standard deviation of 0. The new account balance defined benefit plans have variability (SD) close to 20% relative to the median value of the projected life annuity. Finally, defined contribution plans that are subject to market variability have variability (SD) in the range of 100% [1].

REFERENCES

[1] Bishop, J., Schumacher, P., Olinsky, A. (2011). *Retirement Plan Simulation.* Northeast Decision Sciences Institute Conference.

[2] Green, L. (2003, October 29). *What is a pension equity plan?*. Retrieved from http://www.bls.gov/opub/cwc/cm20031016ar01p1.htm

[3] IRS. (2009, September 5). *Choosing a retirement plan: defined benefit plan*. Retrieved from http://www.irs.gov/retirement/article/0,,id=108950,00.html

[4] Lowman, T. (2000, July 7). *Actuarial aspects of cash balance plans*. Retrieved from http://www.soa.org/files/pdf/actuarial_aspects.pdf

[5] McGill, D, Brown, K, Haley, J, & Schieber, S. (2005). *Fundamentals of private pensions*. Oxford, NY: Oxford University Press.

[6] Watson Wyatt. (2003, September). *Court rules that cash balance plan and pep are age discriminatory*. http://www.watsonwyatt.com/us/pubs/Insider/showarticle.asp?ArticleID=11812

[7] McCourt, S. (2006). *Defined benefit and defined contribution plans: a history, market overview and comparative analysis.* Retrieved October 25, 2009, from http://www.ifebp.org/PDF/webexclusive/06feb.pdf

[8] Employee Benefit Research Institute. (2006, March). *Hybrid retirement plans: the retirement income system continues to evolve*. Retrieved from http://www.ebri.org/publications/ib/index.cfm?fa=ibDisp&content_id=80

[9] Benefit Systems, Inc, Initials. (2008). *Cash balance plans*. Retrieved from http://www.benefitsystems.com/newsinfo/newsfiles/cashbalance.pdf

USING DATA ENVELOPMENT ANALYSIS FRAMEWORK TO EVALUATE MUTUAL FUND EFFICIENCY

D.K. Malhotra, Philadelphia University, (215) 951-2813, MalhotraD@philau.edu
Rashmi Malhotra, St. Joseph's University, (610) 660-3497, rmalhotr@sju.edu

ABSTRACT

The study shows the areas in which inefficient mutual funds are lagging behind and how they can improve their performance to bring them at par with other efficient mutual funds. Using data envelopment analysis approach, this study compares the relative efficiency of aggressive growth mutual funds. We study the efficiency of mutual funds in terms of risk-adjusted performance, load, 12b-1 plan, and expense ratios. Using the DEA methodology, we calculated an efficiency score for 189 funds on a scale of 1 to 100 by maximizing twelve month total return and minimizing beta, standard deviation, load, 12b-1 charges, and expense ratios.

Keywords: Mutual Funds, Mutual Fund Expense Ratios, Data Envelopment Analysis

INTRODUCTION

Usually, in the selection of a mutual fund, investors consider historic performance, risk, the investment objective of the fund, and/or the investment manager's style. However, there is no evidence to support positive performance persistence as being more useful than prior performance as a fund selection tool. Studies by Jensen (1968), McDonald (1974), and Crenshaw (1977) do not find evidence of superior performance by open-end mutual funds. Kon and Jen (1979), Chang and Lewellen (1984), and Lee and Rahman (1990) find limited success by fund managers who practice market timing and selectivity.

In recent years, academic studies and the popular press have stressed the importance of examining a fund's expense structure, because expenses directly affect fund returns. Studies by Malkiel (1995), Malhotra and McLeod (1997), and Livingston and O'Neal (1996, 1998) highlight the significance of expenses in selecting an open-end fund. However, we cannot simply select funds on the basis of expense ratios alone, because academic studies also show that there is no relation between expense ratios and performance of a fund.

The need is to select efficient mutual funds based on investment objective, performance, risk, and expense ratio of a fund. In this study, we use data envelopment analysis (DEA) models to separate efficient mutual funds from the inefficient funds. DEA approach will help us evaluate efficiency of a mutual fund on a relative basis instead of absolute performance measurement as given by traditional performance measurement measures. Also, we will be able to include the cost of owning a mutual fund share in the form of a fund's expense ratio, load charges, 12b-1 charges as an input variable in addition to fund's objective, return and risk as measured by beta of the fund and standard deviation of the fund.

Mutual funds have of course been popular investment vehicles. The Investment Company Institute publishes historical statistics about mutual funds. According to these statistics, the first U. S. mutual fund was established in 1924. Since that time, the number of mutual funds increased to 8,545 at the end of 2010 with assets under management of $13.1 trillion. These numbers includes equity, bond, money market, and hybrid funds. The number of U.S. households owning mutual fund shares reached 51.6 million in 2010, which is nearly one-half the total (44% of the U.S. households). Therefore, it is extremely important that investors should be able to distinguish between efficient and inefficient mutual funds. Our study will also help mutual fund manager benchmark their funds against other funds in the same fund family as well as against the competition and help improve performance of the fund.

The rest of the study is organized along the following lines. In section II, we discuss previous studies on fund performance as well as on the use of data envelopment analysis models in financial analysis. Section III provides the model used in this study. Section IV discusses the data source and methodology used in this study. In section V, we provide an empirical analysis of our results. Section VI summarizes and concludes our study.

LITERATURE REVIEW

Numerous studies have analyzed the operating efficiency of firms using data envelopment analysis models. Hung, Lu, and Wang (2010) explore the *operating efficiency*, the scale efficiency targets, and the variability of *DEA* efficiency estimates of Asian container ports. Joo, Min, Kwon, and Kwon (2010) use data envelopment analysis to assess the *operating efficiency* of specialty coffee retailers from the perspective of socially responsible global sourcing. They evaluate the impact of socially responsible sourcing on the *operating efficiencies* of specialty coffee retailers before and after implementing fair-trade practices. Their study also compares the *operating efficiencies* of fair-trade coffee retailers to those of non-fair-trade coffee retailers. Hung and Lu (2008) study applies the Data Envelopment Analysis (*DEA*) approach with the classical radial measure, non-radial efficiency measure and efficiency achievement measure, respectively, combining multiple outputs and inputs to measure the magnitude of performance differences between the IC firms. Shimshak and Lenard (2007) present a Two-Model approach for including quality measures in *DEA* studies. This approach allows decision-makers to evaluate two models simultaneously, one measuring operational efficiency and the second measuring quality efficiency. This new method selects only DMUs that are efficient in both operational and quality measures to be members of the benchmark set. Their study demonstrates the Two-Model *DEA* approach using data from the nursing home industry. Lu, Yang, Hsiao, and Lin (2007) study uses the CCR model of Data Envelopment Analysis (*DEA*) and the slack variable analysis to evaluate the *operating efficiency* of the domestic banks in Taiwan from 1998 to 2004. Using data from the Annual Survey of Hospitals compiled by the Department of Health in Taiwan for years 1994 through 1997, Chang, Chang, Das, and Li (2004) use Data Envelopment Analysis (*DEA*) to evaluate the impact of a National Health Insurance (NHI) Program on the *operating efficiency* of district hospitals in Taiwan. Anderson, Fok, Springer, and Webb (2002) measure the technical efficiency and economies of scale for real estate investment trusts (REIT) by employing data envelopment analysis (*DEA*). Using data from the National Association of Real Estate Investment Trusts (NAREITs) for the years 1992-1996, they report that REITs are technically inefficient, and the inefficiencies are a result of both poor input utilization and failure to operate at constant returns to scale. Golany, Roll, and Ryback (1994) study applies data envelopment analysis (*DEA*) for measuring and evaluating the *operating efficiency* of power plants in the Israeli Electric Corporation is discussed.

A large number of studies have examined mutual fund performance using data envelopment analysis approach. Chehade (1998) uses production models based on DEA methodology to evaluate the performance of Canadian mutual funds by computing efficiency scores. Basso and Funari (2001, 2003) use DEA methodology to develop mutual fund performance index. Anderson, Brockman, Giannikos, and McLeod. (2004) apply DEA models to evaluate real estate mutual funds.

In this paper, we extend previous studies by illustrating the use of DEA models to benchmark the performance of mutual funds in terms of risk-adjusted performance as well as expenses. No previous study has benchmarked mutual funds in terms of return and expenses.

MODEL

The Data Envelopment Analysis Model:
The Data Envelopment Analysis (DEA) (Charnes et al., 1978) is a widely used optimization-based technique that measures the relative performance of decision making units that are characterized by a multiple objectives and/or multiple inputs structure. Data envelopment analysis is a technique used to assess the comparative efficiency of homogenous operating units such as schools, hospitals, utility companies, sales outlets, prisons, and military operations. More recently, it has been applied to banks (Haslem, Scheraga, & Bedingfield, 1999) and mutual funds (Haslem & Scheraga, 2003; Galagedera & Silvapulle, 2002; McMullen & Strong, 1998; Murthi, Choi, & Desai, 1997). 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.

The DEA methodology measures the performance efficiency of organization units called Decision-Making Units (DMUs). This technique aims to measure how efficiently a DMU uses the resources available to generate a set of outputs. The performance of DMUs is assessed in DEA using the concept of efficiency or productivity defined as a ratio of total outputs to total inputs. Efficiencies estimated using DEA are relative, that is, relative to the best performing DMU or DMUs (if multiple DMUs are the most efficient). The most efficient DMU is assigned an efficiency score of unity or 100 percent, and the performance of other DMUs vary between 0 and 100 percent relative to the best performance.

DATA AND METHODOLOGY

Our data is taken from Morningstar Principia. The data set is for March 2011. Our annual data for each fund includes investment objective, twelve month total return, fund's beta, fund's three-year annualized standard deviation, front-end load, deferred load (percent), 12b-1 plans cost (percent), and fund's audited expense ratio (percent). We choose mutual funds with prospectus objective of aggressive growth. There are a total of 189 mutual funds for which the data is available.

EMPIRICAL ANALYSIS

Each of the mutual funds is a homogenous unit, and we can apply the DEA methodology to asses a comparative performance of these funds. The study evaluates the efficiency of the funds that maximize the twelve month total return and minimize a mutual fund's load, 12b-1 plan charges, and expense ratios. Using the DEA methodology, we calculate an efficiency score for 189 funds on a scale of 1 to 100. We analyzed and computed the efficiency of the funds with data for the March, 2011. Table 1 illustrates the efficiency scores and the rankings of the 189 funds as of March, 2011. As illustrated in table 1, VALIC Company II Agrsv Growth Lifestyle, Transamerica Asset Allc Interm Horizon, Delaware Pooled Core Focus Fixed Income, First Investors Special Situations A LW, BlackRock Aggressive Gr Prepared Instl, Wasatch Long/Short, Invesco Dynamics Y, and Perkins Discovery are 100% efficient., and rest of the funds are less than 100% efficient. JPMorgan Growth Advantage Sel, MassMutual Select Aggressive Growth Y, American Century Vista A, American Century Vista A Load Waived, First Investors Special Situations B, MFS Aggressive Growth Allocation 529A, MFS Aggressive Growth Allocation 529A LW, Maxim Aggressive Profile II, and Alger Growth Opportunities C are 99% efficient, and so on. Similarly, SunAmerica Focused Gr B and SunAmerica Focused Gr C are 72% efficient. Midas Special is least efficient at 68% efficiency level at 182nd rank. Figure 1 illustrates the efficiency factor of 189 funds. The pareto-efficient funds on the efficiency frontier are 100% efficient, and the inefficient funds, below the efficiency frontier, are less than 100% efficient..

We present the score in percentage value varying between 0% and 100%. We find that efficiency of VALIC Company II Agrsv Growth Lifestyle, Transamerica Asset Allc Interm Horizon, Delaware Pooled Core Focus Fixed Income, First Investors Special Situations A LW, BlackRock Aggressive Gr Prepared Instl, Wasatch Long/Short, Invesco Dynamics Y, and Perkins Discovery is 100%. On the other hand, rest of the funds rank from 2 to 182 in the order of decreasing efficiency. Figure 2 illustrates the efficiency frontier formed by 100% efficient mutual funds. All the less efficient funds lie below the efficiency frontier. This means that the observed level of total 12 month return of JP Morgan Growth Advantage is .99 times the maximum output level that JP Morgan Growth Advantage can secure with its current beta (3 year), standard deviation (3 year), audited expense ratio, front load, deferred load, and 12b-1 current value. The same rationale applies to the rest of the funds. As 180 funds are inefficient relative to VALIC Company II Agrsv Growth Lifestyle, Transamerica Asset Allc Interm Horizon, Delaware Pooled Core Focus Fixed Income, First Investors Special Situations A LW, BlackRock Aggressive Gr Prepared Instl, Wasatch Long/Short, Invesco Dynamics Y, and Perkins Discovery in Mrach, 2011; the next step is to identify the efficient peer group or funds whose operating practices can serve as a benchmark to improve the performance of these funds. Table 2 illustrates the peer group for the inefficient countries.

As shown in the table, Delaware Pooled Core Focus Fixed Income, Wasatch Long/Short, and First Investors Special Situations A LW serve as peer for Western Asset Absolute Return I. In addition, Western Asset Absolute Return I is more comparable to Delaware Pooled Core Focus Fixed Income (weight 72%), less comparable to its distant peer First Investors Special Situations A LW (27%), and even less comparable to Wasatch Long/Short (1%). Thus, Delaware Pooled Core Focus Fixed Income should scale up its current beta (3 year), standard deviation (3 year), audited expense ratio, front load, deferred load, and 12b-1 current value factors to make them comparable with Delaware Pooled Core

Focus Fixed Income. Similarly, Reaves Select Research Instl has First Investors Special Situations A LW (83%) as the closest peer that it should emulate, VALIC Company II Agrsv Growth Lifestyle (9%) as its next distant peer, Transamerica Asset Allc Interm Horizon (7%) as its far distant peer, and Wasatch Long/Short (1%) as its farthest distant peer. Similarly, we can use the most highly weighted peers of all other inefficient funds to assess their relative efficiency performance and make decisions regarding what factors to change in comparison to the closest peers. Although, we list the efficiency score of Invesco Dynamics Y and Perkins Discovery as 100%, their efficiency level is 1.00129 and 1.00191 respectively, so these funds are not exactly on the efficiency frontier but very near to the frontier and therefore have corresponding peer funds.

First Investors Special Situations A LW serves as the closest peer, and the second closest peer for all the inefficient funds. Similarly, BlackRock Aggressive Gr Prepared Instl serves as the most immediate or immediate peer for most of the inefficient funds. On the other hand, Delaware Pooled Core Focus Fixed Income is the immediate peer or the distant peer for some of the inefficient funds. Similarly, Wasatch Long/Short is the distant peer for some of the inefficient funds. VALIC Company II Agrsv Growth Lifestyle is the distant peer or the farther distant peer for a small number of the inefficient funds. Finally, Transamerica Asset Allc Interm Horizon is farthest distant peer for three of the inefficient funds. Therefore, First Investors Special Situations A LW is the most efficient fund among all the funds under study as not only is the First Investors Special Situations A LW 100% efficient, but it also serves as the role model for all funds. Similarly, BlackRock Aggressive Gr Prepared Instl is also the next most efficient fund among the group of funds under study. The efficient peer funds have a similar mix of input-output levels to that of the corresponding inefficient funds, but at more absolute levels. The efficient funds generally have lower input levels relative to the fund in question. The features of efficient peer funds make them very useful as role models inefficient funds can emulate to improve their performance. Furthermore, First Investors Special Situations A LW is used as an efficient peer to all Pareto-inefficient funds, so its frequency of use as an efficient-peer, expressed as a percentage of the number of pareto-inefficient countries, is 100%. BlackRock Aggressive Gr Prepared Instl is an efficient peer to 103 of inefficient funds with a frequency rate of 58%. Wasatch Long/Short and Delaware Pooled Core Focus Fixed Income is an efficient peer to 30 funds with net percentage of 16%. In addition, VALIC Company II Agrsv Growth Lifestyle has the peer efficiency frequencies of 13%. Finally, Transamerica Asset Allc Interm Horizon has a benchmarking factor of 2%. Thus, we have enhanced confidence that First Investors Special Situations A LW is the most outperforming mutual fund followed by BlackRock Aggressive Gr Prepared Instl as they outperform all the other funds. Furthermore, these funds are more likely to be a better role model for less efficient funds to emulate because their operating practices and environment match more closely those of the bulk of inefficient funds. Table 3 displays the benchmarking factor and the hit percentage of efficient fund.

After calculating the efficiency of a fund 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 fund within the possible set of input-output levels. The DEA efficiency measure tells us whether or not a fund can improve its performance relative to the set of funds to which it is being compared. Therefore, after maximizing the output efficiency, the next stage involves calculating the optimal set of slack values with 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 fund (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 fund cannot be expanded jointly or individually without raising its input level. Further, its input level cannot be lowered given its output levels. Thus, the funds are pareto-efficient with technical output efficiency of 1. If the fund 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 fund A in question, but it offers in excess of the fund A on the output corresponding to the positive slack. Thus, fund A is not Pareto-efficient, but with radial efficiency of 1 as its output cannot be expanded jointly. Finally, if the fund A is not efficient (<100%) or the efficiency factor is greater than 1, then the fund 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 fund 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 fund A is not reflected in its efficiency measure because the additional output does not apply across all output dimensions. Table 5 illustrates the slack values identified in the next stage of the DEA analysis. The slack variables for 100% efficient funds are zero. Therefore, VALIC Company II Agrsv Growth Lifestyle, Transamerica Asset All c Interm Horizon, Delaware Pooled Core Focus Fixed Income, First Investors Special Situations A LW, BlackRock Aggressive Gr Prepared Instl, Wasatch Long/Short, Invesco Dynamics Y, and Perkins Discovery is 100%. Pareto-efficient as the DEA model has been unable to identify some feasible production point which can improve on some other input or output level. On the other hand, for Western Asset Absolute Return I, decreasing the input level of Audited Expense Ratio by 0.2010 units, will enable the fund to outperform.. Western Asset Absolute Return I can follow Delaware Pooled Core Focus Fixed Income and First Investors Special Situations A LW as its role model and emulate their policies. Similarly, Western Asset Absolute Return FI can reduce its Audited Expense Ratio level by 0.3113 units and 12b-1 Current value by 0.1676 units while maintaining efficient levels equivalent to that of its peers—Delaware Pooled Core Focus Fixed Income and First Investors Special Situations A LW. On the same lines, we can find the slack factors for the underperforming funds Table 4 illustrates the slack values of the relevant factors for inefficient countries.

SUMMARY AND CONCLUSIONS

Using data envelopment analysis approach, this study compares the relative efficiency of aggressive growth mutual funds. We study the efficiency of mutual funds in terms of risk-adjusted performance, load, 12b-1 plan, and expense ratios. Using the DEA methodology, we calculated an efficiency score for 189 funds on a scale of 1 to 100 by maximizing twelve month total return and minimizing beta, standard deviation, load, 12b-1 charges, and expense ratios. There are seven aggressive growth mutual funds that are 100 percent efficient. Rests of the mutual funds are less than 100% efficient. The study also shows the areas in which inefficient mutual funds are lagging behind and how they can improve their performance to bring them at par with other efficient mutual funds.

The data envelopment analysis is a powerful technique for performance measurement. The major strength of DEA is its objectivity. DEA identifies efficiency ratings based on numeric data as opposed to subjective human judgment and opinion. In addition, DEA can handle multiple input and outputs measured in different units. Also, unlike statistical methods of performance analysis, DEA is non-parametric, and does not assume a functional form relating inputs and outputs.

However, as with any other study, this study using DEA has certain limitations (Ramanathan, 2003). The application of DEA involves solving a separate linear program for each DMU. Thus, the use of DEA can be computationally intensive. In addition, as DMU is an extreme point technique, errors in measurement can cause significant problems. DEA efficiencies are very sensitive to even small errors, thus making sensitivity analysis an important component of post-DEA procedure. Also, as DEA is a non-parametric technique, statistical hypothesis tests are difficult to apply. Therefore, further extension of this study would be to perform principal component analysis of the all the DEA model combinations. Furthermore, we can also use logistic regression to test the validity of the results.

TABLES, FIGURES, & REFERENCES

Tables, figures, references, and full paper available upon request from the authors.

COMPARISON OF CURRENCY CO-MOVEMENT BEFORE AND AFTER OCTOBER 2008

M. E. Malliaris, Loyola University Chicago, 1 E. Pearson, Chicago, IL,
mmallia@luc.edu, 312-915-7064
A.G. Malliaris, Loyola University Chicago, 1 E. Person, Chicago, IL,
tmallia@luc.edu, 312-915-6063

ABSTRACT

This paper considers foreign currency markets and discusses how currencies can be used as investment, hedging and speculative instruments. These three financial activities take place because certain fundamentals relationships exist between and among currencies that persist despite the surrounding uncertainty. To uncover such relationships and to examine their robustness over time and across various currencies, Apriori association analysis is performed on two sets of data, prior to the September 2008 financial crisis and after. This paper focuses specifically on the directional movement of eight major currencies in these two time periods and investigates whether there are any Apriori rules common to both periods. We then look at the performance of several of these rules in each time period, on both training and validation sets.

INTRODUCTION

The daily volume of currency transactions in currency futures, forwards, swaps and options dominates all other types of trading volumes. This volume is driven by globalization that includes trade and foreign direct investments, global portfolio diversification, global hedging and speculation and global banking, among other factors. Changes in information and communication technologies have accelerated the processing and transmission of data and ideas to a level far beyond our capabilities of a decade or two ago.

Global financial imbalances have also played an important role in the volume and behavior of currencies. Substantial trading surpluses in China, Germany and Japan have been transferred in countries with balance of payments deficits such as the U.S. Global banks have facilitated the transfer of funds from subsidiaries in countries with large saving to countries with low saving and attractive investment opportunities. Also, the establishment of sovereign funds by surplus countries has contributed to the movement of currencies.

Increased financial globalization has offered important opportunities for portfolio diversification. There is a wide spectrum of financial risks that include firm specific risks, industry wide risks and country risks. Several macroeconomic policies such as monetary policy, fiscal policy and level of regulation are included in country risks, among several other factors. The growth in the size and complexity of international financial markets has been one of the most striking aspects of the world economy over the last decade. Lane and Milesi-Ferretti [6] document the increase in gross holdings of cross-country bond and equities for a large group of countries. They describe this as a process of financial globalization. Economists and policy makers have

speculated on the implications of financial globalization for the design of monetary policy. Most central banks now either explicitly or implicitly follow a policy of inflation targeting. Under this policy, price stability, appropriately defined, is the principal goal of monetary policy.

Campbell et al. [3] considered an equity investor who chooses fixed currency weights to minimize the unconditional variance of her portfolio. Such an investor wishes to hold currencies that are negatively correlated with equities. Their first novel result is that at one extreme, the Australian dollar and the Canadian dollar are positively correlated with local-currency returns on equity markets around the world, including their own domestic markets. At the other extreme, the euro and the Swiss franc are negatively correlated with world stock returns and their own domestic stock returns. The Japanese yen, the British pound, and the US dollar fall in the middle, with the yen and the pound more similar to the Australian and Canadian dollars, and the US dollar more similar to the euro and the Swiss franc.

When considering currencies in pairs, Campbell et al. found that risk-minimizing equity investors should short those currencies that are more positively correlated with equity returns and should hold long positions in those currencies that are more negatively correlated with returns. When considering seven currencies as a group, they found that optimal currency positions tend to be long on the US dollar, the Swiss franc, and the euro, and short on the other currencies. A long position in the US-Canadian exchange rate is a particularly effective hedge against equity risk.

This brief discussion points to the conclusion that certain global currencies have received great significance the past few decades. When we also consider the creation of the euro this significance becomes even greater since several important national currencies such as the German marc, the French franc, the Italian lira and several others were replaced by the euro. This global significance translates into a search for the pricing of these currencies and for understanding of their co-movement. The challenge becomes even greater since currencies are priced one in terms of another. One may view the issue of pricing currencies as a comparison of all economic and financial fundamentals between two nations.

Instead of pricing currencies, in this paper we study the directional behavior of leading currency prices over two periods – prior to and following the crash of October 2008. It is well known that the Global Financial Crisis of 2007-09 has impacted all financial markets very pronouncedly. The initial crash of the credit markets with the collapse of the values of mortgage backed securities and housing values quickly expanded to global equity, bond and currency markets. Global banks experienced substantial losses in their portfolios of asset backed securities, the Fed moved aggressively to reduce interest rates and establish several credit facilities to supply much needed liquidity. Foreign currency markets experienced substantial volatilities.

The Global Financial Crisis started in August of 2007 but no one at that time had any accurate assessment of how the crisis would evolve. After Lehman Brothers filed for Chapter 11 bankruptcy protection on September 15, 2008, the intensity of the crisis amplified. Our main goal is to find whether or not there are any relationships among currency movements that have remained stable prior to and following the crash, using October 1 as our dividing date. If price movements are not random but follow patterns, such information may allow economists to better

evaluate the overall impact of leading currencies in the global economy. They can also be used as a basis for formulating speculative, hedging and carry trade strategies.

Mandelbrot and Hudson [8] give a detailed description of the presence of non-linear determinism in financial markets. Empirical evidence of chaotic dynamics in financial data such as stock market indexes, foreign currencies, macroeconomic time series and several others have been performed by various researchers, Kyrtsou and Vorlow [5] recently, and in much more detail earlier by Brock, Scheinkman and LeBaron [9] and Brock and Malliaris [2]. However, this paper will study co-movements in currency markets before and after October 2008, and ask "Is there stability in patterns of directional co-movement of major currencies before and after this time period?

DATA AND METHODOLOGY

Data for eight currency prices, the Australian dollar, the Japanese yen, the Euro, the Swiss franc, the British Pound, the Canadian dollar, the Mexican peso, and the Brazilian real, was retrieved from the Bank of Canada nominal noon exchange rates, which are published each business day at about 12:30 ET and were downloaded from http://www.bankofcanada.ca/rates/exchange/10-year-converter/ Each original number is the amount of the currency equal to one US dollar on that day at that time. The data covered a time period beginning in November 2005, and ending in September 2011.

During the week that began on October 6, 2008, the Dow Jones fell over 18% and the S&P 500 fell more than 20%. This crash was followed by declines in other markets around the world. The month of October 2008 is thus a dividing time in our data set. We removed this month, dividing the data set into two distinct pieces: Before (October 2008) and After. Each data set contains over 700 days of data. The Before and After sets were further subdivided into training and validation sets with the validation set being the last 252 days of each set. The validation sets thus occur entirely after the training sets and are completely disjoint. This type of disjoint, temporally following, and lengthy validation set is the most difficult for a model to perform well on and will thus be a very good judge of the rules stability. The training sets were used to generate rules of directional movement for the currencies. The validation sets were used to judge the robustness of rules on entirely new data. The four sets are named Before Training, Before Validation, After Training, and After Validation. The beginning and ending dates for each set are shown in Table 1.

Table 1. Data Sets

Set Name	Begin Date	End Date	# Rows
Before Training	11/1/2005	9/28/2007	480
Before Validation	10/1/2007	9/30/2008	252
After Training	11/3/2008	9/17/2010	471
After Validation	9/20/2010	9/19/2011	252

The data was originally downloaded as numeric values. These were converted into category-type data that represented the direction of movement of the respective currency relative to the US dollar. The values used were Up, Even, and Down.

Though association analysis originated with the study of market baskets to see which items people purchased at the same time, it has been generalized to look at questions of what occurs together. It is often in an exploratory way to discover interesting relationships in the data that may be analyzed further. For an in-depth discussion of association analysis techniques, see Hand et al [4], or Berry and Linoff [1]. The methodology used to generate rules on these currencies was Apriori Association Analysis, run in IBM's SPSS Modeler data mining package.
Association analysis generates a set of rules of the form IF A THEN B where variables used in the modeling process may occur either after the IF or after the THEN. The set of rules that is generated also depend on the user-supplied minimum values of support and confidence. Support refers to the percent of times that some combination of inputs (also called antecedents) occurs in the data set. This forms the IF part of the statement. When the antecedent combination does occur, confidence reflects the percent of time that the output, or consequent, is also true. This is the confidence of the THEN part of the rule. There are several major association analysis techniques, for example, Apriori, Generalized Rule Induction (GRI), and Carma. These vary in the way they search for interesting rules within a large, and generally sparse, data set. In this problem, support and confidence were set to 7% and 65%, respectively. That is, for any rule to be generated, the IF part must occur in the training data set at least 7% of the time, and when the IF part is there, the THEN part must be true at least 65% of the time.

RESULTS

Using the settings for support and confidence detailed above, the Before training set generated 2635 rules. The After training set generated 2643 rules. While there are many ways to look at the results from these training runs, we will focus on those rules that occurred in both of the training sets. These are the rules that have remained stable on both sides of the October 2008 crash.

Of these two rule sets, 79 rules had identical antecedents and consequents. That is, 79 common rules of the form IF [antecedent] THEN [consequent] occur in both the Before and After training sets. From these 79, one rule with the greatest confidence in the Before training set for each possible market direction represented was selected for further analysis. There were eleven such rules. Not every market and direction combination generated a rule common to both sets. No rules were generated using the market direction EVEN since there are not enough days were this occurs for a rule to be created. These common rules are shown in Table 2.

The antecedents contain multiple conditions, all of which must be true for the rule to be applicable. For example, Rule 1 states that if the Japanese yen was Down (relative to the US dollar) and the Mexican peso was Down and the Brazilian real was Down then the Australian dollar was Down.

Table 2. Rules common to both Before and After training sets

Rule	Antecedent	Consequent
1	DirJpy = Down and DirMex = Down and DirBrz = Down	DirAus = Down
2	DirMex = Up and DirSws = Down and DirAus = Down	DirEur = Down
3	DirBrz = Up and DirSws = Up and DirMex = Down	DirEur = Up
4	DirMex = Up and DirSws = Down	DirJpy = Down
5	DirAus = Up and DirSws = Up and DirMex = Down	DirJpy = Up
6	DirSws = Down and DirCan = Down and DirAus = Down	DirPnd = Down
7	DirAus = Up and DirCan = Up and DirSws = Up and DirMex = Down	DirPnd = Up
8	DirBrz = Up and DirMex = Up and DirEur = Down	DirSws = Down
9	DirEur = Up and DirMex = Down	DirSws = Up
10	DirSws = Up and DirCan = Down and DirAus = Down	DirMex = Down
11	DirBrz = Up and DirCan = Up and DirSws = Up	DirMex = Up

Each of these eleven rules occurred in both the Before and After training sets, but with different values for support and confidence. Though the minimum values for these are set at run-time, the Modeler package calculates the actual values of support and confidence for each rule that occurred in the data set used for training. These are shown in Table 3. Notice that all confidence values are above 70%. This means that, when the If part of the rule was satisfied, the THEN part was true at least 70% of the time. However, a much harder test is the comparison on each validation set.

Table 3. Values of support and confidence for each training set

Rule	Before Support	Before Confidence	After Support	After Confidence
1	15.63	84.00	17.20	79.01
2	15.83	92.11	7.86	81.08
3	7.29	85.71	9.13	90.70
4	23.75	74.56	17.41	70.73
5	14.17	88.24	9.98	70.21
6	26.25	84.13	28.03	82.58
7	7.50	94.44	7.01	78.79
8	15.21	89.04	8.07	84.21
9	23.13	91.89	19.32	82.42
10	9.58	76.09	11.04	75.00
11	14.17	75.00	23.57	75.68

The validation set data is from the year following the training set in each of the cases. These results on the respective Validation sets are shown in Table 4.

Table 4. Values of support and confidence for each validation set.

Rule	Before Support	Before Confidence	After Support	After Confidence
1	16.67	80.95	25.50	93.75
2	14.29	86.11	9.16	56.52
3	9.52	87.50	6.37	87.50
4	25.40	93.75	24.30	85.25
5	13.49	67.65	7.57	78.95
6	23.81	80.00	24.70	70.97
7	8.73	95.45	2.39	66.67
8	12.30	100.00	7.17	77.78
9	24.21	88.52	15.94	60.00
10	11.51	89.66	14.34	86.11
11	19.05	68.75	15.54	89.74

An easier way, perhaps to look at these numbers is in how they change from the training to the validation sets. This is summarized in Table 5.

Table 5. Validation values minus Training values.

Rule	Before Support	Before Confidence	After Support	After Confidence
1	1.04	-3.05	8.30	14.74
2	-1.55	-5.99	1.31	-24.56
3	2.23	1.79	-2.76	-3.20
4	1.65	19.19	6.89	14.51
5	-0.67	-20.59	-2.41	8.73
6	-2.44	-4.13	-3.32	-11.61
7	1.23	1.01	-4.62	-12.12
8	-2.91	10.96	-0.90	-6.43
9	1.08	-3.37	-3.38	-22.42
10	1.92	13.57	3.30	11.11
11	4.88	-6.25	-8.03	14.07

Here we can see that, in the rules generated with the Before training set, while the rules still had a usably strong percent of occurrence, the confidence on the validation set dropped in one direction for each of the represented markets. In contrast, the After training and validation sets differences show that all the drops in confidence occurred with consequents of the European currencies, the Euro, the British pound, and the Swiss franc. Rules with consequents relating to the Australian dollar, the Japanese yen, and the Mexican peso remained positively robust and grew stronger.

CONCLUSIONS

This paper considers foreign currency markets and discusses how currencies can be used as investment, hedging and speculative instruments. These three financial activities take place

because certain fundamentals relationships exist between and among currencies that persist despite the surrounding uncertainty. To uncover such relationships and to examine their robustness over time and across various currencies, an a priori association analysis is performed on two sets of data, prior to the September 2008 financial crisis and after.

There are 11 rules that this paper identifies as most appropriate for further analysis. All 11 rules, if analyzed one at a time, confirm stable relationships among currencies that would allow hedging and speculative activities. For example, if certain currencies decline in a given day then certain other currency will also decline. Also, if certain currencies rise in a given day another currency will also rise. Between all three currencies rising or all three declining and as a consequence also observing a certain currency also rising or also declining, we have several mixed cases. For example if the Swiss franc and Australian dollar decrease while the Mexican Peso increases the rule suggests that the euro will also decline.

Overall, it is confirmed in these rules that the euro, the Japanese yen and the euro move in the same direction most of the times; also it is confirmed that the Mexican peso and Brazilian real move together although the Canadian dollar also influences the Mexican peso. The British pound is influenced by, and in turn it influences, the Australian dollar.

In conclusion, these rules demonstrate that stable and robust relationships exist among groups or pairs of currencies that in turn form the fundamentals for global banking and investment, hedging and speculative activities.

REFERENCES

[1] Berry, M. and Linoff, G. Data Mining Techniques, Second Edition, Indianapolis, IN: Wiley Publishing Inc., 2004.

[2] Brock, W and Malliaris, A. Differential Equations, Stability, and Chaos In Dynamic Economics, Amsterdam, Netherlands: Elsevier Science, 1989.

[3] Campbell, JY, Medeiros KS-de, Viceira LM. *Global Currency Hedging*. Journal of Finance, 2010, 65(1):87-122.

[4] Hand, D., Mannila, H., and Smyth, P. Principles of Data Mining, Cambridge, MA: The MIT Press, 2001.

[5] Kyrtsou C and Vorlow C. *Modelling Nonlinear Comovements Between Time Series*. Journal of Macroeconomics, 2009, 31(1): 200–211.

[6] Lane, P., and Milesi-Ferretti, G.M. *The External Wealth of Nations: Measures of Foreign Assets and Liabilities for Industrial and Developing Countries.* Journal of International Economics, 2001, 55, pp. 263-94.

[7] Lane, P., and Milesi-Ferretti, G.M. *The External Wealth of Nations Mark II*. IMF Working Paper, No 06-69, 2006.

[8] Mandelbrot, B. and Hudson, R. The (Mis)Behavior of Markets. New York, NY: Basic Books, 2004.

[9] Scheinkman, Jose A and LeBaron, B. *Nonlinear Dynamics and Stock Returns*. Journal of Business, University of Chicago Press, 1989, 62(3), pages 311-37.

AN EMPIRICAL ANALYSIS OF COST EFFICIENCIES IN THE INDIAN BANKING INDUSTRY

D.K. Malhotra, Philadelphia University, (215) 951-2813, MalhotraD@philau.edu
Raymond Poteau, Philadelphia University, (215) 951-2811, PoteauR@philau.edu

ABSTRACT

This study evaluates cost efficiencies in the Indian banking industry for the period 2007 to 2011. Empirical results show that total cost as well as operating cost of a bank is positively related to the size of the bank defined in terms of assets and deposits. This study shows that total cost and operating cost increases less than proportionately to increase in assets and deposits since 2009, which points to economies of scale with reference to assets and deposits. Results also show that ownership structure of a bank plays a role in determining economies of scale in the Indian banking industry. Cost increases in the private sector banks have been less than proportionate to increases in assets and deposits throughout the sample period.

Keywords: Economies of Scale, Banking, Translog Cost Function

INTRODUCTION

The banking industry in India has come a long way from the nationalization of its banks in 1969 to the liberalization of the financial system since 1991. During the reform process that started in 1991, the banking sector was opened up with the objective of improving the efficiency of the banking system in India through increased competition from private and foreign banks. With this view, the government initiated the process of removing interest rate controls. The government also introduced capital adequacy requirements and other safety norms to ensure a sound banking system. The objective is to strengthen banking supervision and increase competition through licensing of private banks and foreign banks. The ultimate goal is to integrate Indian banks into the global financial system.

Although the amount of assets and deposits managed by the Indian banking industry has increased several folds, empirical research on economies of scale and cost efficiencies in the Indian banking industry are still limited. The issue of economies of scale and cost efficiencies in the banking industry is important for several reasons.

Firstly, the Indian economy is the second fastest growing economy in the world after China. Since these nations represent an engine of growth for the world economy, a large amount of capital is flowing to the stock markets of these nations. Banks, as financial intermediaries, are playing a crucial role by bringing enhanced liquidity and promoting market efficiency by facilitating smooth transfer of funds between borrowers and lenders that will promote capital mobility among nations. A sound and efficient banking system is essential for a smooth integration of Indian financial markets with the rest of the world, because banks play a crucial role in facilitating transfer of funds between borrowers and lenders. Therefore, the size of the banking system is bound to grow further.
Secondly, previous studies show that a country's financial sector influences future economic growth. The banking sector is the most important part of the financial markets. If the Indian banking sector is sound and efficient, it will have a positive impact on India's growth.

Thirdly, we examine the cost efficiencies of Indian commercial banks during the period 2007 to 2011. This time period covers before crisis and during crisis time. Findings from this study will highlight the reaction of cost efficiencies in the Indian banking sector during and after the global financial crisis.

This study will also help the banking industry as well as regulatory agencies such as the Reserve Bank of India (RBI) get a better understanding of the impact of growth of a bank on bank's expenses and its benefits to bank shareholders in the form of higher returns and increased shareholder wealth.

We distinguish our study from previous studies in two ways. Firstly, no study has looked at economies of scale over a longer period of time. We study economies of scale over a period of six years from 2007 to 2011 to investigate the consistency of economies of scale over a longer period of time.

Secondly, this study covers the period of the global financial meltdown and we evaluate the cost efficiencies of the Indian banks during and after this meltdown.

This paper has six sections. Section II provides a discussion of previous research related to this topic. Section III briefly describes our data. Section IV describes our methodology and gives information on banks used in this study. Section V presents our empirical results. Section VI gives our conclusions.

DATA AND METHODOLOGY

The data sample consists of 20 state owned banks and 15 private banks. Data covers the fiscal year ending March 31st 2007 to March 31st 2011. Table 1 provides a summary statistics of the variables used in this study. Table 1 shows that total expenses for the banks as a whole gone up by 110 percent and operating expenses have increased by 115 percent from 2007 to 2011. During the same period, total assets have increased by 113 percent and total deposits have gone up by 117 percent. Thus, there is less than proportionate increase in total expenses relative to assets and deposits. Operating expenses show a higher percentage increase relative to increase in assets, but slightly less increase relative to increase in deposits.

As shown in Table 1, during the period 2007-2011, percentage increase in total expenses and operating expenses is higher for public sector banks at 119% and 125%, respectively. In fact, total assets and total deposits of public sector banks have also gone up by approximately the same percentage over the period of 2007 and 2011.

Private sector banks, on the other hand, show a significantly lower percentage increase in total expenses and operating expenses at 86% and 90%, respectively. The assets and deposits of private sector banks have gone up by 94% and 89%, respectively during the same period.

Our methodology involves estimation of the coefficients of a translog cost function to determine which factors contribute to economies of scale and their degree of contribution. We then estimate cost elasticity with respect to the amount of assets using the first derivative of the translog cost function. Cost elasticity is estimated for the total sample for each year and for subsets of the annual samples. The subsets are created according to ownership of the bank.

In order to investigate economies of scale in banks, we use a two-part methodology. The first part is an estimation of coefficients for a translog cost function to determine which factors contribute to economies of scale and the extent to which they contribute for each of the five years in the period 2007 to 2011. .

In financial economics, the translog model is the most pervasive approach for investigating economies of scale. The translog cost model implicitly assumes a U-shaped average cost function. It is used here because it allows economies of scale to vary with level of bank assets.

The estimation of scale economies with a translog cost function requires cost and output measures. For commercial banks, the output can be defined in two ways. It can be total assets under management and it can also be total deposits that the bank is able to capture in the competitive market. Total cost of each bank is defined as the total expenses of operating its banks, including its management fee. A bank's total operating expense is modeled as a function of total assets and control variables that affect level of expenses.

$$\text{Ln COST} = ß_0 + ß_1 \ln \text{DEPOSITS} + ½\, ß_2(\ln \text{DEPOSITS})^2 + e \qquad (1a)$$

$$\text{Ln OPERATING COST} = ß_0 + ß_1 \ln \text{DEPOSITS} + ½\, ß_2(\ln \text{DEPOSITS})^2 + e \qquad (1b)$$

Equation 2 is the translog cost function with respect to deposits of a bank.

$$\text{Ln COST} = ß_0 + ß_1 \ln \text{ASSETS} + ½\, ß_2(\ln \text{ASSETS})^2 + e \qquad (2a)$$

$$\text{Ln OPERATING COST} = ß_0 + ß_1 \ln \text{ASSETS} + ½\, ß_2(\ln \text{ASSETS})^2 + e \qquad (2b)$$

In the translog function, COST is the rupee amount of a bank's total expenses expenses, OPERATING COST is the cost of operations of a bank, ASSETS is total bank assets, and DEPOSITS includes total deposits of bank.

The most common measure of operating efficiency in economies of scale studies is the elasticity of cost with respect to the output. When the rate of increase in output exceeds the rate of increase in cost in an industry, economies of scale characterize that industry. For the banking industry, cost elasticity with respect to assets can be used to evaluate the existence and extent of economies of scale. It is measured by percentage change in cost associated with a percentage change in bank assets. We calculate this elasticity by taking the first derivative of the translog cost function (Equation 1) with respect to assets.

The result is Equation 3.

$$\frac{\partial(\ln Y)}{\partial(\ln X)} = \beta_1 + \beta_2(\ln X) \qquad (3)$$

Where Y represents total expenses of a bank for equation 1a and 2a and it also represents operating expenses of a bank for equations 1b and 2b. X represents total assets of a bank for equation 1 and total deposits of a bank for equation 2.

If cost elasticity is less than one, bank's expenses increase less than proportionately with changes in bank assets. This implies that economies of scale exist. If the elasticity is greater than one, we can infer that diseconomies of scale exist.

To investigate the existence of economies of scale, we estimate the scale economy measure for each observation and then average across observations to derive the group scale economy measure. The cost elasticity is found for each observation (bank). Then an average across observations is computed to obtain the group average elasticity.

We estimate cost elasticities for the total group of banks in each annual sample. We also estimate elasticities for groups within each annual sample where the groupings are according to ownership—private sector or public sector banks.

EMPIRICAL ANALYSIS

Regression results for total annual samples of banks are shown in Table 2. Some of the results are as expected, while others are not. Model 1 in Table 2 shows that our model in equation 1a explains 97% to 98% of the total expenses. The coefficient on natural logarithm of deposits is positive, which means that, on an average, an increase in the deposits of the bank leads to higher expenses for all the banks in the sample. Cost elasticity equals 1.005 in 2007 and is not statistically different from 1, which means that, on an average, expenses increased in the same proportion as the increase in assets in 2007. However, beginning with 2008, cost elasticity is declining and in 2011, the cost elasticity is 0.91 and it is statistically significant, which means that as the deposits of a bank increase, total expenses increase less than proportionately and, on an average, there are economies of scale for all the banks in the sample.

Model 2 also explains 98% to 99% of the total expenses of a bank. Model 2 in Table 2 shows that in four out of five years, size of the bank as measured by assets is positively related to total expenses. As bank size grows, total expenses also increase and the coefficient on natural logarithm of assets is statistically significant in explaining the total expenses in four out five years in the sample. Cost elasticity equals 1.03 in 2007 and is statistically significant, which points to diseconomies of scale for that year. Since 2008, cost elasticity is less than one and is statistically significant, which means economies of scale exist for the Indian commercial banks as whole from 2008 to 2011.

Table 3 shows the regression results for the translog cost function for operating expenses of banks with respect to total assets and total deposits. On an average, our model 1 explains 89% to 99% of the operating expenses of a bank with respect to total deposits of a bank. Only in 2008, the relationship between total deposits and operating expenses is positive and statistically significant. Cost elasticity of operating expenses with respect to deposits is greater than one at 1.16 and is highly statistically significant in 2007. Since 2008, cost elasticity is lower than one and is consistently declining. As banks add more deposits, operating expenses are rising less than proportionately, which points to economies of scale in operating expenses with respect to total deposits.

Model 2 in Table 3 also shows that when we analyze operating expenses with respect to total assets, our model explains 92% to 95% of operating expenses. Cost elasticity is less than one and statistically significant through the sample period except for the year 2008. Thus, there are economies of scale in operations when the size of the bank increases in terms of assets.

Part A of Table 4 shows cost elasticity of total expenses and operating expenses with respect to total deposits when we analyze economies of scale by bank ownership. On an average, cost elasticity of total expenses with respect to total deposits is less than one for private sector banks and is statistically significant for each year in the sample. Total expenses of private sector banks do not increase in the same proportion as the increase in deposits of private sector banks, which points to economies of scale throughout the sample period.

On the other hand, cost elasticity of total expenses with respect to total deposits for public sector banks is less than one in 2009 and 2011 only. In 2007 and 2008, cost elasticity is higher than one and is statistically significant, which points to diseconomies of scale for public sector banks during these two years. In 2010, cost elasticity is not statistically different from one for public sector banks.

Cost elasticity of operating expenses with respect to total deposits for private sector banks is less than one and statistically significant for the all the years except for the year 2007. Public sector banks show cost elasticity higher than one through the sample period except for the year 2011 when they have a cost elasticity that is not statistically different from one. Therefore, on an average, operating expenses of public sector banks increase more than proportionately with respect to increase in deposits, which points to diseconomies of scale.

Table 4 also shows that private sector banks also show economies of scale in their total expenses and operating expenses with respect to total assets of the banks with cost elasticity less than one throughout the sample period. Therefore, on an average, if the size of the private sector banks has gone up in terms of assets, total expenses and operating expenses have increased less than proportionately. Public sector banks show consistent economies of scale in total expenses with respect to total assets since 2008. In operating expenses, public sector banks have started reaping economies of scale since 2010 with cost elasticity less than one during 2010 and 2011.

In all the above cases, cost elasticity for private sector banks is less than the cost elasticity for public sector banks, which means private sector banks have relatively higher economies of scale in comparison to their public sector counterparts.

Table 5 ranks banks on the basis of economies of scale with the highest rank being given to a bank that has the maximum economies of scale and is considered to be the most cost effective. As show in Table 5, Dhanalakshmi Bank is the most efficient banks in terms of lower cost with respect to deposits as well as assets for the years 2007 to 2009. Dhanalakshmi bank has a cost elasticity of total expenses with respect to total deposits of 0.93 in 2007 and it declines to 0.81 in 2009. In 2010 and 2011, Development Credit Bank rises to the top spot with a cost elasticity of 0.84 and 0.66 and Dhanalakshmi slips to second place in the rankings. Both these banks are private sector banks.

Table 5 also shows that the State Bank of India (SBI), a public sector bank, is the most inefficient bank with a cost elasticity of more than one throughout the sample period. In fact, cost elasticity for SBI has increased from 1.07 in 2009 to 1.17 in 2011.

Table 6 ranks banks on the basis of cost elasticity of operating expenses with respect to total deposits. Again, Dhanalakshmi is the most efficient bank from 2007 to 2009 and Development Credit Bank is the most efficient with lower operating expense increase with every increase in deposits in 2010 and 2011. SBI continues to be the most inefficient bank, but in operating expenses cost elasticity has gone down from 1.41 in 2007 to 1.17 in 2011. Table 5 and Table 6 shows that the second most inefficient bank is ICICI bank, which is a private sector bank.

Tables 7 and 8 show similar rankings of banks on the basis of cost elasticity of total expenses and operating expenses with respect to total assets. SBI continues to be at the lowest rank in terms of cost efficiencies followed by ICICI bank.

SUMMARY AND CONCLUSIONS

India started the process of economic reforms and opened up its banking sector in 1991. The goal was to improve the efficiency of the Indian banking sector with an aim to integrate into the global financial sector. This study analyzed the cost efficiencies of 35 Indian commercial banks over the period 2007 to 2011. We used translog cost function to evaluate economies of scale in the Indian banking sector. Translog cost function was estimated for total cost and for operating cost

with respect to total deposits and total assets of 35 banks. Estimates of cost elasticity showed that Indian commercial banks are reaping economies of scale with a cost elasticity of less than one. This study showed that total cost and operating cost increases less than proportionately to increase in assets and deposits since 2009, which points to economies of scale with reference to assets and deposits. Results also show that ownership structure of a bank plays a role in determining economies of scale in the Indian banking industry. Cost increases in the private sector banks have been less than proportionate to increases in assets and deposits throughout the sample period.

TABLES, FIGURES, & REFERENCES

Tables, figures, references, and full paper available upon request from the authors

Simulation Modeling of Patient Workflow in Hospital Clinics

Sung J. Shim
Seton Hall University, USA
sung.shim@shu.edu

Arun Kumar
RMIT University, Australia
a.kumar@rmit.edu.au

Jianxin (Roger) Jiao
Georgia Tech Savannah, USA
roger.jiao@me.gatech.edu

Organizations reengineer their business processes to contain costs, improve efficiency, and stay competitive in the marketplace. With escalating healthcare costs, hospitals also seek ways to contain costs and provide quality healthcare services. Hospitals have traditionally emphasized breakthroughs in healthcare procedures and technology to stay competitive. As competition among hospitals continues to intensify, however, patients may perceive little difference in healthcare procedures and technology used by different hospitals. Consequently, hospitals come to understand that process reengineering could be a better solution to achieve competitive advantage. Just as many businesses successfully reduce costs and gain competitive advantage by reengineering their business processes, hospitals can reengineer the way certain healthcare processes are carried out to achieve efficiency and cost containment. Computer simulation, which has proven successful in improving various business processes, can also be an effective tool in searching for more efficient processes in hospitals.

This paper describes a case study undertaken at outpatient clinics of Tan Tock Seng Hospital in Singapore. In order to improve patient workflow in outpatient clinics, the hospital management considers implementing some changes in patient workflow along with new systems for patient appointment scheduling and online billing. Using computer simulation, the study models the patient workflow in outpatient clinics and assesses patient turnaround times in the patient workflow. Then, it evaluates the effects of implementing the changes along with the systems for patient appointment scheduling and online billing on patient turnaround times in the patient workflow. The results of the study will prove helpful to those considering reengineering and improving patient workflow or other similar processes in hospitals.

FACTORY EXPANSION AT A CONTRACT MANUFACTURER OF PHARMACEUTICAL INGREDIENTS

Christian Grandzol, Bloomsburg University, 400 East Second St. Bloomsburg PA, 17815, cgrandzo@bloomu.edu, 570-389-4521
Ed Pitingolo, Bloomsburg University, 400 East Second St. Bloomsburg PA, 17815, epintingo@bloomu.edu, 570-389-4519

ABSTRACT

This teaching case traces the conversion of Winding Creek Pharmaceuticals to a contract manufacturer in 2008. For decades, the business had been part of a "Big Pharma" company that determined the plant's production levels, requirements, and other key decisions, but structural changes in the pharmaceutical industry led to the sale of Winding Creek to a relatively small business that did not have manufacturing or pharmaceuticals experience. To support the growth ambitions of its new owner, Winding Creek evaluated an opportunity to substantially increase factory capacity.
Keywords: Theory of Constraints, Production, Contract Manufacturer, Pharmaceutical Industry

INTRODUCTION

This case was inspired by a field trip taken by students and faculty to Winding Creek Pharmaceuticals in April, 2011. At the time, Winding Creek informed that they had recently been sold back to a "Big Pharma" company after a 3 year stint on their own as a contract manufacturer. Company leaders discussed differences between operating in a captive supply chain and competing in the market, expectations established for Winding Creek when it was bought, significant changes occurring in the pharmaceutical industry, and several decisions the business recently faced.

A specific decision that caught the interest of the authors and the students was the decision to expand the capacity of one factory at the plant. The expansion would consume considerable resources just as access to financing was becoming more difficult due to a global recession. With long cycle times to develop sales and a myriad of other issues, there was debate whether expanding the factory was essential and if mid-2008 was an ideal time to begin the 1-year expansion. Winding Creek had an interesting story to share with students and after multiple in-depth interviews and plant visits, we wrote the Winding Creek case.

This case fills a gap in the management case literature particularly in the area of Operations Management and Supply Chain by affording students an opportunity to engage in decision making, analyze management's actions, and apply their disciplinary knowledge in the context of a factory expansion. The case examines the Theory of Constraints, production issues, utilization rates, overhead dilution, and the pharmaceutical industry. The case is targeted for use in upper-level, undergraduate business courses such as Operations Management.

OVERVIEW OF THE CASE

The Winding Creek plant had 50 years of experience offering on-time, in-spec active pharmaceutical ingredients (API) to its Big Pharma owner, Limerick Stone Pharmaceuticals. Due to a variety of reasons, Limerick Stone and other Big Pharma companies found themselves with substantial excess capacity in the U.S. Limerick Stone decided to sell its plant in late 2007 to ALPHA, a relatively small, government services organization that had neither pharmaceutical nor manufacturing experience. The plant was re-named Winding Creek Pharmaceuticals and it set out with unbridled hopes to compete as a contract manufacturer of API and other services.

The case delves into numerous issues the new contract manufacturer faced. For example, competing as a stand-alone required a change in culture, the ability to accurately price projects, and significant customer development. Winding Creek found it difficult to grow sales as quickly as it needed due to intense competition from established global contract manufacturers, a contracting economy, a slowing overall pharmaceutical industry, and structural changes in the industry.

To support the growth in sales required to sustain the business, Winding Creek decided to add filter dryer capacity in its main chemistry block, Factory 29. The dryers were the bottleneck in the process—the extra capacity would allow the running of multiple products simultaneously and could potentially double or triple output. The benefits of adding the filter dryers made the expansion an "essential move."

Rex Quinn, the plant manager, agreed with the expansion, but was not sure about its implementation. Was the spring of 2008 the correct time to start the expansion and spend $10 million executing it? Projections and assumptions made during the time of the decision were proving unrealistic or faulty. What would making an error on Factory 29 (the organization's major strategic project) mean to the young contract manufacturer? Would a "stop" from him, a plant manager, matter anyway?

Why do parents change their pediatricians and

when do the opt for private-but-expensive children hospitals?

Dr. V. Miligkos, MD, PhD
Staff Pediatrician, Pediatrics Clinic, Mitera S.A. Greece

M. V. Miligkos, MD, PhD Candidate
Tufts Medical Centre, Tufts University, Boston, Massachusetts

Dr. Konstantinos Nikolopoulos*, PhD
Professor of Decision Sciences, Bangor Business School, U.K.

Dr. Sally Sambrook
Professor of Human Resources Development, Bangor Business School, U.K.

***Corresponding Author: k.nikolopoulos@bangor.ac.uk; mob: 0044 7981 332913

Why do parents change their pediatricians and when do they opt for private-but-expensive children hospitals?

Extended Abstract

In this study, following up earlier research presented in NEDSI 2011 on "How Do Parents Select Their Pediatrician?" we investigate the process by which parents *change* their children's primary health care providers, i.e. their Pediatricians. Furthermore we are particularly interested in when parents are prepared to reside to private – but at the same time expensive – hospitals for theirs children's healthcare, especially in emergency situations.

For this purpose, a close-ended questionnaire has been administered to 200 parents in Greece with an approximately 100% response rate due to nature of the collection process. A big part of the responses came from a group of parents that had already changed their pediatrician at least once at the time of the survey, while a smaller part has visited a private children's hospital in a case of emergency. The parents' pediatrician at the time of the survey, was involved in the information collection process. One complementary subgroup was also considered; the very Pediatricians themselves, aiming to obtain some insight on the potentially discrepancies between what specialists think is important and what actually matters. An additional small part of the sample consisted of mothers in the first four days of their first childbirth – although we merged this sample with the original sample of parents.

Our empirical results demonstrate that for all the groups, the priorities concerned primarily the pediatricians' ability to avoid mistakes in the diagnosis and treatment of their children as well as the cost itself, and nevertheless the overall communication skills of the doctor. As far as the willingness to visit a private children's hospital in a case of emergency we found that – not surprisingly – it is significantly associated with the cost of access to such pediatrics facilities.

Finally we offer a Human Resources (HR) perspective, through a discussion of the implications of our findings for HR Management i.e. how private childcare clinics should recruit their specialists, and HR Development i.e. how private childcare clinics should train their specialists.

ISSUES IN ENROLLMENT OF MINORITY STUDENTS IN MIS

Sathasivam Mathiyalakan
Accounting & MIS, School of Business & Economics, Winston Salem State University
Email: mathiyalakansa@wssu.edu

George Heilman
Accounting & MIS, School of Business & Economics, Winston Salem State University
Email: heilmang@wssu.edu

ABSTRACT

Despite the slowing US economy and outsourcing, studies have shown that MIS careers are the fastest growing and tend to be higher paying among undergraduate disciplines. However, there continues to exist an enrollment problem, i.e. not enough students are choosing to major in MIS. Prior studies have shown that the participation of minorities and women in the technology sectors such as MIS and Computer science have been limited and thereby exacerbating the problem of "technology diversity" in the workplace. In this study, we examine issues that play a part in minority students' decision to major in MIS.

Keywords: Minority Students, MIS enrollment, Student Perception, Technology

The Intersection of Technology Use in the Classroom and Sustainability

Cynthia L. Knott

G. Steube

Hongqiang (Mason) Yang

Northeast Decision Science Institute (NEDSI)

Newport, RI

March 21 to 23, 2012

Abstract

The use of technology in universities and colleges is an issue of interest and speculation. One issue related to technology use in the classroom is sustainability of resources that support the technology. This paper explores faculty perceptions about technology use and sustainability in an east coast university. This university has initiated a new program that has been charged with the objective of creating and maintaining a sustainability program. The program is still being developed but a few of the key goals are to look at recycling campus wide, printing costs in the computer labs and library and exploring what the faculty perceptions are about using technology in the classroom. This paper focuses on the last goal and in order to explore this objective a survey was administered to the Schools of Business, Health Professions, Arts & Sciences and the Library.

The research question addressed in this paper is the relationship between the use of technology in the classroom and the course discipline of the faculty teaching the class. The faculty participants in the survey included four of its academic schools: Business, Health Professions, Arts & Sciences and the Library. Consequently there are four different faculty affiliations based on their school. The research questions related to school assignment presented in the paper include the following:

- Are there significant differences in the use of technology based on the school in which the faculty member is associated?
- Do technology-driven programs in the schools of business and library science tend to perceive the use technology in the classroom differently than other schools?
- Are differences in the school of affiliation reflected in faculty views of importance of technology to the learning process?

- Are differences in the school of affiliation reflected in faculty perceptions in the use of technology devices including the desktop computers, iPad/Tablets, Laptops, Smartphones, or E-Readers?
- Do faculty affiliations with schools impact their view of the importance moving toward the use of electronic documents?
- Do faculty affiliations impact whether technology devices are viewed as distractions?
- Are sustainability enthusiasts also technology enthusiasts?

The development of technology can have an impact of teaching. New technologies encourage educators to leverage these developments in the classroom (Klopfer, Osterweil, Groff, & Haas, 2009). Technology also changes the way instruction is delivered by offering educators effective ways to reach different types of learners and assess student understanding through a wider array of options. It can also alter the relationship between instructor and student; effective technology integration moves instructors into the roles of adviser, content expert, and coach (Edutopia, 2011). The use of technology over the past 40 years has been investigated by Tamim, Bernard, Borokhovski, Abrami, and Schmid (2011). The role of technology in the classroom has accelerated over time and will continue to be a factor in classrooms of the future.

At the same time that the growth of technology has occurred there has been a growing interest in sustainability. Sustainability is based on the simple principle that survival and well-being depends, either directly or indirectly, on the natural environment (United States Environmental Protection Agency, n.d.). Colleges and universities have also turned their attention to addressing sustainability. The Association for the Advancement of Sustainability in Higher Education (AASHE) (2011) holds conferences, workshops, and webinars on a regular basis to promote sustainability. Two major objectives of the AASHE are to make sustainable practices the norm within higher education and to the efforts of higher education institutions to integrate sustainability into teaching, research, operations, and public engagement. The goal of this paper is to examine the relationship between technology use in the classroom and sustainability in higher education.

Procedure

A survey was administered to faculty at a small private university on the east coast. Four of the university's academic schools participated in the survey; the Schools of Business, Health

Professions, Arts & Sciences, and Library were included. A copy of the survey questions is available in Appendix A. The purpose of the research was to gather information about the use of technology in the classroom and sustainability. The subject university has initiated a new program that has been charged with the objective of creating and maintaining a sustainability program. Some of the goals of this program include: campus-wide recycling, reduction of hard copy output, and to explore faculty perceptions about using technology in the classroom to augment sustainability efforts.

Population

The target population for this research is full time faculty at the university. This research sampled faculty in four of its schools of instruction. The sample size was 108. It is believed that the results of the sampling process used are generalizable to the target population of faculty at this subject university and at other similar institutions and reflects attitudes and behaviors about the use of technology in the classroom and sustainability.

Research Questions

The research questions addressed in this paper focused on the relationship between the use of technology in the classroom and sustainability. The faculty participants in the survey included four of its academic schools: Business, Health Professions, Arts & Sciences and the Library; the School of Education and Human Services did not participate in this study. Consequently there are four different faculty affiliations based on their school. The research questions related to school assignment presented in the paper include the following:

- Are there significant differences in the use of technology based on the school in which the faculty member is associated?

- Do technology-driven programs in the schools of business and library science tend to perceive the use technology in the classroom differently than other schools?
- Are differences in the school of affiliation reflected in faculty views of importance of technology to the learning process?
- Are differences in the school of affiliation reflected in faculty perceptions in the use of technology devices including the desktop computers, iPad/Tablets, Laptops, Smartphones, or E-Readers?
- Do faculty affiliations with schools impact their view of the importance moving toward electronic documents?
- Do faculty affiliations impact whether technology devices are viewed as distractions?
- Are sustainability enthusiasts also technology enthusiasts?

Survey Instrument

Table 1 identifies the sources of the analysis from the survey questions. In this table each question and a brief caption is presented to identify the relationships that were examined. Several variables were recoded to develop categorical groups for the analysis; these groups and their labels are in Table 2. Tables 1 and 2 connect the questions in the survey to the variables that will be examined in this research. The results of this analysis are presented in the following section.

Table 1

Questions and Captions

Question	Caption
1	School Affiliation
2	Faculty Rank
3	Years of Teaching Experience
4	Importance of Technology in Teaching
5	Permit Electronic Devices for Note Taking
6	Importance of Technology in the Learning Process
7	Importance of Desktop Computers in the Classroom
8	Importance of iPad/Tablet in the Classroom
9	Importance of Laptop/Netbook in the Classroom
10	Importance of Cell Phones in the Classroom
11	Importance of E-Readers in the Classroom
12	Level of Distraction Caused by Electronic Devices
13	Authority for Decisions about Electronic Devices
14	Frequency of Using Handouts in the Classroom
15	Shifting to Electronic Documents

Table 2

Categorical Groups

Question	Category Name (Value)
1. Which school are you working at?	
a. School of Arts/ Sciences	Arts/Science (1)
b. School of Business Administration	Business (2)
c. School of Education	Education (3)
d. School of Health Professions	(4) Health
e. Library and Learning Services	(5) Library
2. What is your faculty status at the Marymount University?	
a. Professor	
b. Associate Professor	
c. Assistance Professor	
d. Term Appointment	
e. Adjunct or Instructor	
3. How many years have you been teaching college students?	
a. More than 10 years	Experienced (1)
b. 5 to 10 years	Experienced (1)
c. 2 to 5 years	Less Experienced (0)
d. Less than 2 years	Less Experienced (0)
4. How important do you think classroom electronics, such as desktop PC, projector, or Smartboard, is to your teaching?	
a. Extremely Important	Highly Important (1)
b. Somewhat Important	Highly Important (1)
c. Important	Important (2)
d. Not very Important	Not Important (3)
d. Not Important	Not Important (3)
5. Do you allow laptops or other electronic devices in your classroom for students to take notes and access files electronically?	
Y	Yes (1)
N	No (0)
6. How important do you think laptops or other electronic devices are to the learning process in the classroom?	
a. Extremely Important	Important (1)
b. Somewhat Important	Important (1)
c. Important	Important (1)
d. Not very Important	Not Important (0)
d. Not Important	Not Important (0)

7. How would you rate the importance of desktop computer in the classroom?

a. Extremely Important	Important (1)
b. Somewhat Important	Important (1)
c. Important	Important (1)
d. Not Important	Not Important (0)

8. How would you rate the importance of iPad/tablets in the classroom?

a. Extremely Important	Important (1)
b. Somewhat Important	Important (1)
c. Important	Important (1)
d. Not Important	Not Important (0)

9. How would you rate the importance of laptop/netbooks in the classroom?

a. Extremely Important	Important (1)
b. Somewhat Important	Important (1)
c. Important	Important (1)
d. Not Important	Not Important (0)

10. How would you rate the importance of cell phones in the classroom?

a. Extremely Important	Important (1)
b. Somewhat Important	Important (1)
c. Important	Important (1)
d. Not Important	Not Important (0)

11. How would you rate the importance of e-readers in the classroom?

a. Extremely Important	Important (1)
b. Somewhat Important	Important (1)
c. Important	Important (1)
d. Not Important	Not Important (0)

12. How distracting are those electronic devices when students use them in the classroom?

a. Extremely Distracting	Distracting (1)
b. Somewhat Distracting	Distracting (1)
c. Distracting	Distracting (1)
d. Not very Distracting	Not Distracting (0)
d. Not Distracting	Not Distracting (0)

14. How frequently do you require students to bring printed handouts to class?

a. Once a week	Frequent (1)
b. A few times a month	Less Frequent (0)
c. A few times a semester	Less Frequent (0)
d. Don’t require at any time	Less Frequent (0)

15. How difficult would it be to make the shift to using electronic documents/information instead of printed handouts in the classroom?	
a. Extremely Difficult	Difficult (1)
b. Somewhat Difficult	Difficult (1)
c. Difficult	Difficult (1)
d. Not very Difficult	Not Difficult (0)
d. Not Difficult	Not Difficult (0)

Survey Discussion

This section discusses the results of the survey based on the based answers to the question. In the following section the results of answering the research questions are presented. Table 3 displays the results for Question 1. The majority of participants were from the School of Arts and Sciences.

Table 3

Question 1

	School			
	Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Arts/Sciences	48	44.4	44.4	44.4
Business	27	25.0	25.0	69.4
Health	25	23.1	23.1	92.6
Library	8	7.4	7.4	100.0
Total	108	100.0	100.0	

Information about faculty rank from Question 2 is displayed in Table 4. Over 50% of the participants in the survey were at the full or associate professor rank.

Table 4

Question 2

	Rank			
	Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Missing Data	4	3.7	3.7	3.7
Professor	27	25.0	25.0	28.7
Associate Professor	29	26.9	26.9	55.6
Assistant Professor	43	39.8	39.8	95.4
Term Appointment	4	3.7	3.7	99.1
Adjunct or Instructor	1	.9	.9	100.0
Total	108	100.0	100.0	

Years of teaching experience as captured in Question 3 are provided in Table 5. The majority of faculty who responded to the survey had over 10 years of teaching experience.

Table 5

Question 3

	Teaching Experience			
	Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
More than 10 years	67	62.0	65.7	65.7
5 to 10 years	21	19.4	20.6	86.3
2 to 5 years	9	8.3	8.8	95.1
Less than 2 years	5	4.6	4.9	100.0
Total	102	94.4	100.0	
Missing Data	6	5.6		
Total	108	100.0		

Table 6 presents the results to Question 4 which identified the importance of laptops and other electronic devices is to the teaching process in the classroom. The majority of faculty (73%) view technology as extremely important to the teaching process in the classroom.

Table 6

Question 4

Importance of Technology to the Teaching Process in the Classroom				
	Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Extremely Important	79	73.1	76.0	76.0
Somewhat Important	13	12.0	12.5	88.5
Important	4	3.7	3.8	92.3
Not very Important	5	4.6	4.8	97.1
Not Important	3	2.8	2.9	100.0
Total	104	96.3	100.0	
Missing Data	4	3.7		
Total	108	100.0		

Question 5 addressed the use of laptops or other electronic devices in the classroom. Table 7 reveals that 93% of the faculty allows the use of an electronic device for note taking or accessing information in the classroom.

Table 7

Question 5

Laptop use in the Classroom					
		Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
	Yes	93	86.1	89.4	89.4
	No	11	10.2	10.6	100.0
	Total	104	96.3	100.0	
Missing	Data	4	3.7		
Total		108	100.0		

The importance of laptops or other devices in the learning process in the classroom is asked in Question 6. Table 8 shows that only 10% of the faculty surveyed considered technology as not important in the learning process.

Table 8

Question 6

Importance of Technology to the Learning Process in the Classroom				
	Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Extremely Important	27	25.0	27.3	27.3
Somewhat Important	36	33.3	36.4	63.6
Important	7	6.5	7.1	70.7
Not Very Important	19	17.6	19.2	89.9
Not Important	10	9.3	10.1	100.0
Total	99	91.7	100.0	
Missing Data	9	8.3		
Total	108	100.0		

Tables 9 thru 13 provide the results for Questions 7 through 12 which assessed five technology devices. The ranking of devices that received the highest scores as extremely important are desktop PC, laptop/notebook, iPad/tablet, e-reader, and cell phone.

Table 9

Question 7

Desktop Importance in the Classroom				
	Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Extremely Important	50	46.3	55.6	55.6
Somewhat Important	11	10.2	12.2	67.8
Important	7	6.5	7.8	75.6
Not Important	22	20.4	24.4	100.0
Total	90	83.3	100.0	
Missing Data	18	16.7		
Total	108	100.0		

Table 10

Question 8

iPad/Table Importance in the Classroom				
	Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Extremely Important	13	12.0	13.4	13.4
Somewhat Important	20	18.5	20.6	34.0
Important	11	10.2	11.3	45.4
Not Important	53	49.1	54.6	100.0
Total	97	89.8	100.0	
Missing Data	11	10.2		
Total	108	100.0		

Table 11

Question 9

Laptop/Netbook Importance in the Classroom				
	Frequency	Percent	Valid Percent	Cumulative Percent
Extremely Important	23	21.3	24.5	24.5
Somewhat Important	25	23.1	26.6	51.1
Important	9	8.3	9.6	60.6
Not Important	37	34.3	39.4	100.0
Total	94	87.0	100.0	
Missing Data	14	13.0		
Total	108	100.0		

Table 12

Question 10

Cell Phone Importance in the Classroom				
	Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Extremely Important	2	1.9	2.1	2.1
Somewhat Important	2	1.9	2.1	4.3
Important	7	6.5	7.4	11.7
Not Important	83	76.9	88.3	100.0
Total	94	87.0	100.0	
Missing Data	14	13.0		
Total	108	100.0		

Table 13

Question 11

E-Reader Importance in the Classroom				
	Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Extremely Important	8	7.4	8.7	8.7
Somewhat Important	16	14.8	17.4	26.1
Important	9	8.3	9.8	35.9
Not Important	59	54.6	64.1	100.0
Total	92	85.2	100.0	
Missing Data	16	14.8		
Total	108	100.0		

The perceived level of distraction arising from the use of electronic devices in the classroom is provided in Table 14. Less than 20% of the respondents felt that technology devices were not very distracting or not distracting.

Table 14

Question 12

	Distraction			
	Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Extremely Distracting	34	31.5	34.7	34.7
Somewhat Distracting	27	25.0	27.6	62.2
Distracting	16	14.8	16.3	78.6
Not very Distracting	14	13.0	14.3	92.9
Not Distracting	7	6.5	7.1	100.0
Total	98	90.7	100.0	
Missing Data	10	9.3		
Total	108	100.0		

Authority for making decisions to allow or ban the use of technology devices in the classroom is presented in Table 15. Almost all participants in the survey agreed that authority for the use of electronic devices in the classroom should be determined by faculty.

Table 15

Question 13

		Authority			
		Frequency	Percent	Valid Percent	Cumulative Percent
	Instructor	99	91.7	96.1	96.1
	University	4	3.7	3.9	100.0
	Total	103	95.4	100.0	
Missing	Data	5	4.6		
Total		108	100.0		

The frequency of using printed handouts in the classroom is displayed in Table 16. About 40% of the faculty surveyed did not require the use of printed handouts at any time in the course.

Table 16

Question 14

Frequency of Using Printed Handouts				
	Frequency	Percent	Valid Percent	Cumulative Percent
Once a week	19	17.6	20.0	20.0
A few times a month	10	9.3	10.5	30.5
A few times a semester	22	20.4	23.2	53.7
Don't require at any time	44	40.7	46.3	100.0
Total	95	88.0	100.0	
Missing Data	13	12.0		
Total	108	100.0		

The difficulty in transitioning to using all electronic documents instead of printed handouts is provided in Table 17. A slight majority (50.9%) of faculty felt that there would be some degree of difficulty in making the transition to electronic documents.

Table 17

Question 15

Shifting to Electronic Documents				
	Frequency	Percent	Valid Percent	Cumulative Percent
Extremely Difficult	21	19.4	20.6	20.6
Somewhat Difficult	26	24.1	25.5	46.1
Difficult	8	7.4	7.8	53.9
Not Very Difficult	22	20.4	21.6	75.5
Not Difficult	25	23.1	24.5	100.0
Total	102	94.4	100.0	
Missing Data	6	5.6		
Total	108	100.0		

Research Questions Discussion

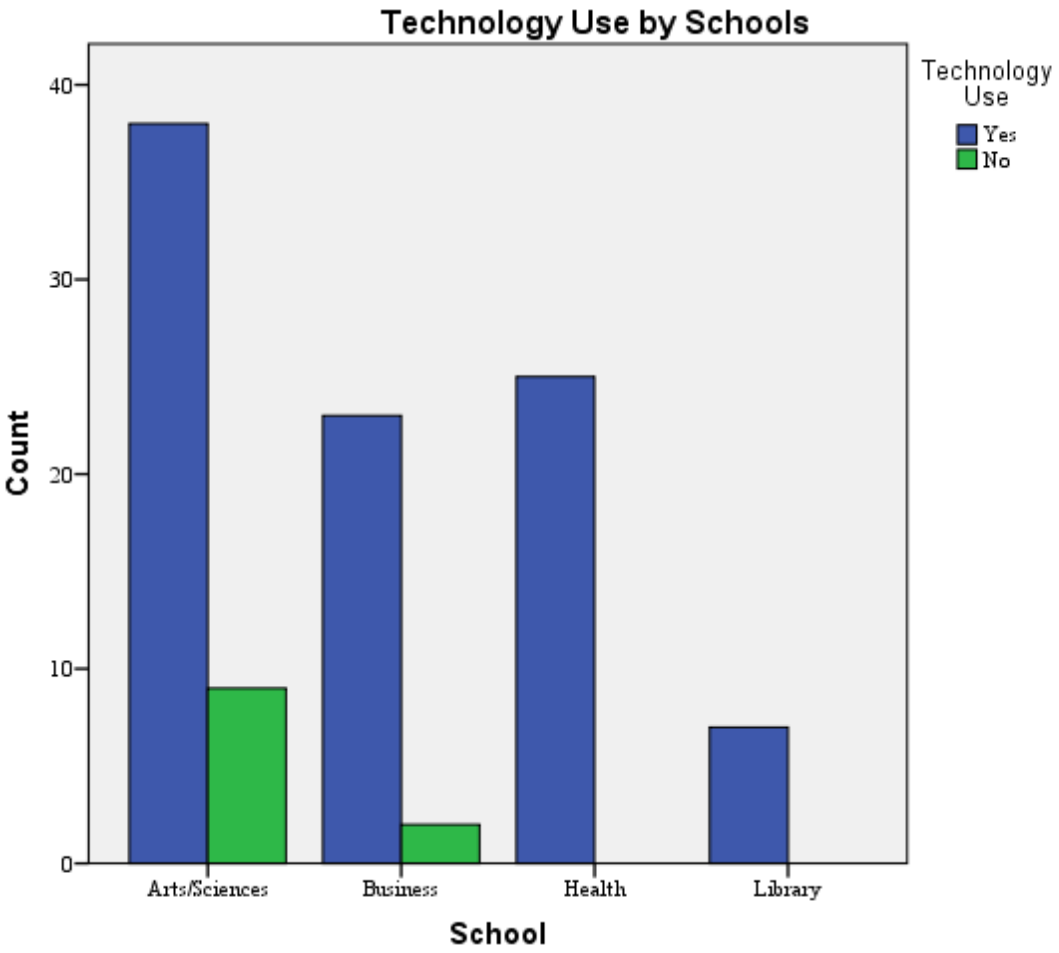
This section discusses the following research questions: (1) Are there significant differences in the use of technology based on the school in which the faculty member is associated? (2) Do technology-driven programs in the schools of business and library science tend to perceive the use technology in the classroom differently than other schools? (3) Are differences in the school of affiliation reflected in faculty views of importance of technology to the learning process? (4) Are differences in the school of affiliation reflected in faculty perceptions in the use of technology devices including the desktop computers, iPad/Tablets, Laptops, Smartphones, or E-Readers? (5) Do faculty affiliations with schools impact their view of the importance moving toward the use of electronic documents? (6) Do faculty affiliations impact whether technology devices are viewed as distractions? (7) Are sustainability enthusiasts also technology enthusiasts?

These questions were analyzed using contingency tables with a chi-square to test the existence of an association and with a phi coefficient to assess the strength of the associations. Phi is a chi-square based measure of association; the chi-square coefficient depends on the strength of the relationship and sample size. Since phi has a known sampling distribution it is possible to compute its standard error and significance (Howell, 2002). For this analysis the strength of the association will be assessed through a rule of thumb which provides a range of values for Phi and verbal assessment. Strong negative and strong positive associations are represented by Phi values between -1.0 to .-7 and .7 to 1.0, respectively. Weak negative and positive associations are between -.7 to .-3 and .3 to .7, respectively. Values of Phi indicating little or no association are between -.3 to .3 (Simon, 2005).

Research Question 1

The first research question examines the relationship between the use of technology in the classroom and the faculty member's school affiliation. Questions 1 and 5 from the survey were investigated to explore the relationship. Figure 1 displays the distribution of technology use in the classroom by academic affiliation. The relationship was not significant at the .05 level χ^2 (3) = 7.612, p = .055; the strength of the association was also low with a Phi coefficient of .271.

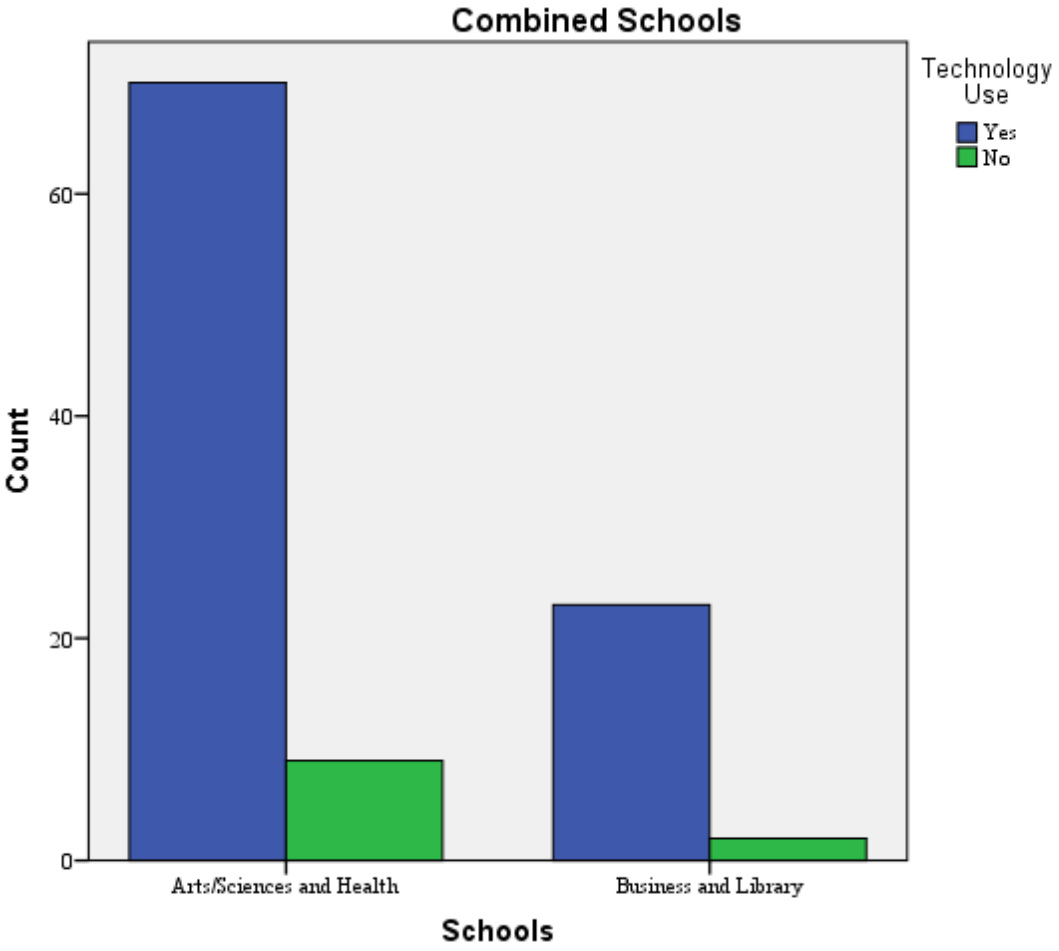
Figure 1



Research Question 2

Research question 2 examined the relationship between the use of technology in programs in the schools of business and library science in the classroom differs from other schools. Questions 1 and 5 were used for this analysis. The data for the schools of business and library science were combined in one group; the data for the schools of arts and sciences and library services was also pooled into a single group. Figure 2 displays the results of this analysis. There was no significant relationship between the combined schools and their use of technology in the classroom, χ^2 (1) = .231, p =.631; the Phi coefficient also indicated the lack of a relationship with a value of -.047.

Figure 2

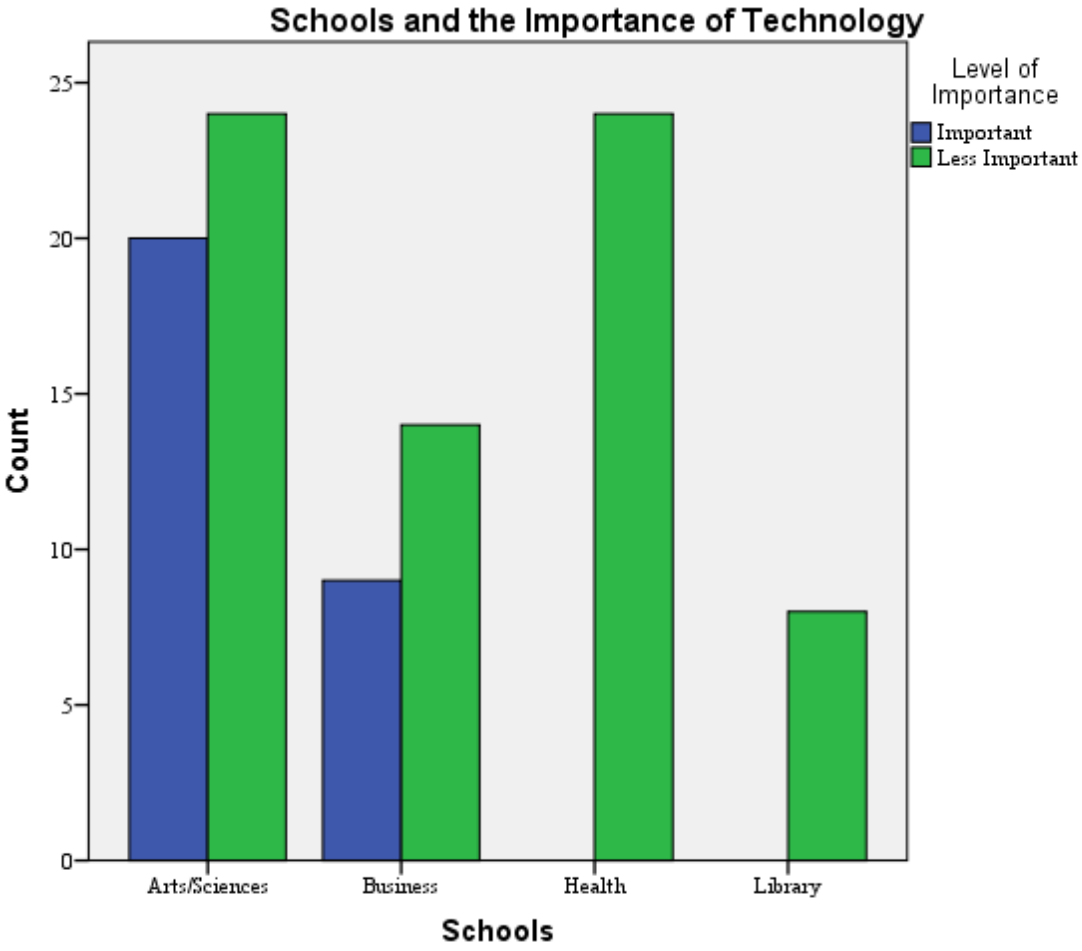


Research Question 3

This question investigated the relationship between school of affiliation and faculty views of the importance of technology to the learning process. Questions 1 and 6 were used to inspect this association. Figure 3 displays the graph for this analysis. In question 6, the first three choices (extremely important, somewhat important, and important) were grouped to create the category of important; responses of not very important and not important were combined into the less important class. At the .05 level of significance there is a relationship the school of affiliation

and the perception of the importance of technology in the classroom to learning, χ^2 (3) = 19.881, p < .0; the Phi coefficient was .448 which indicates a weak association.

Figure 3

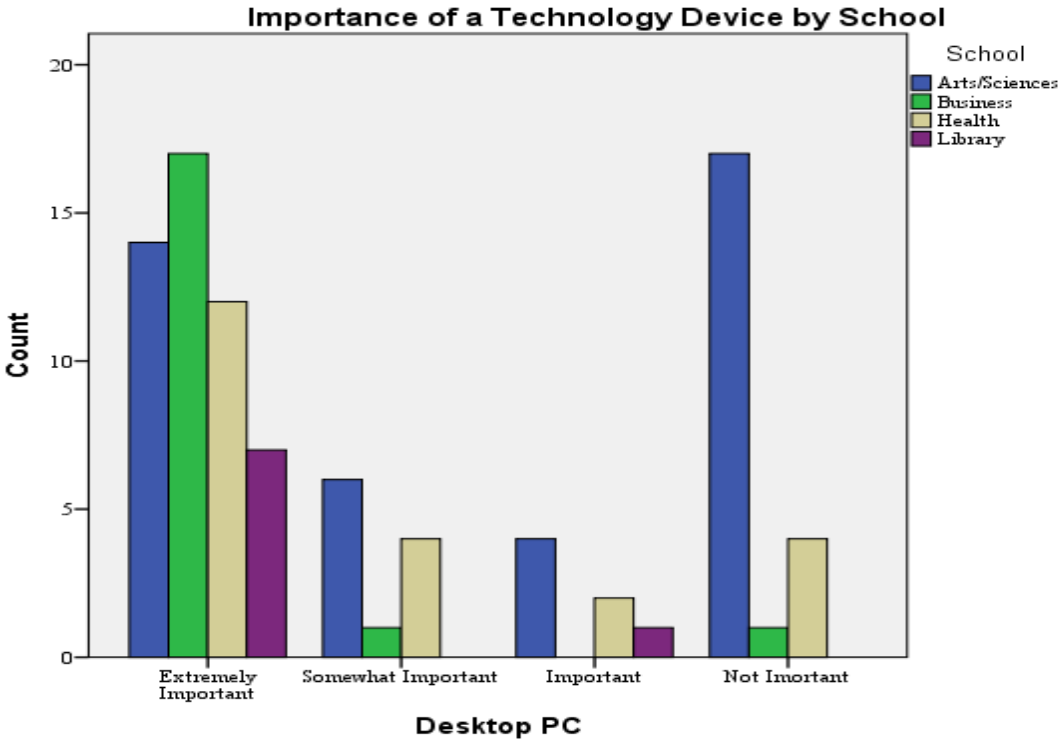


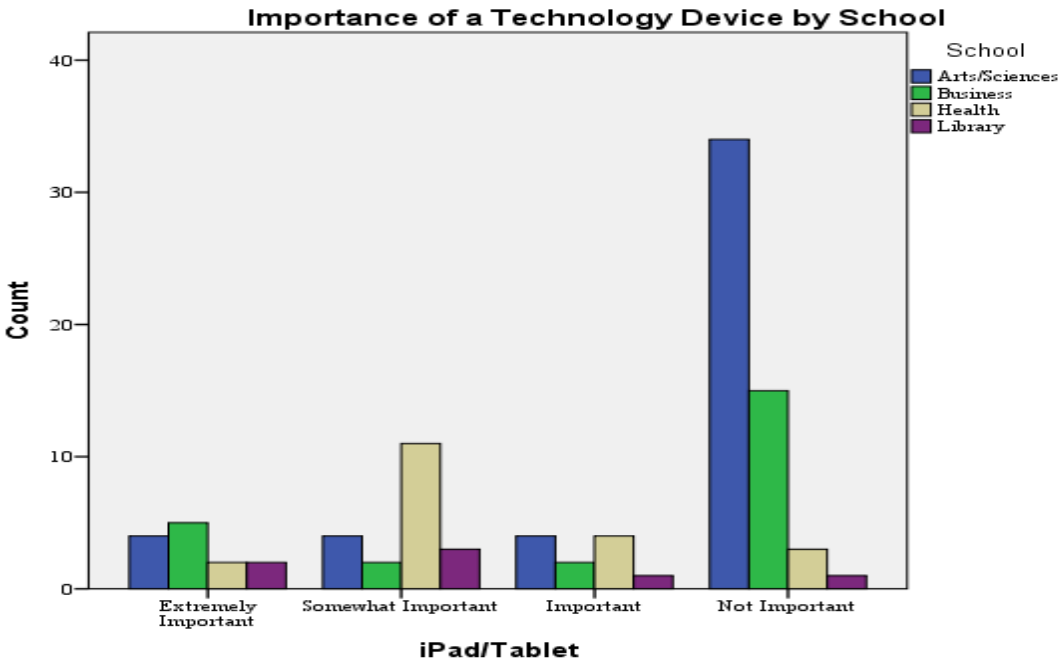
Research Question 4

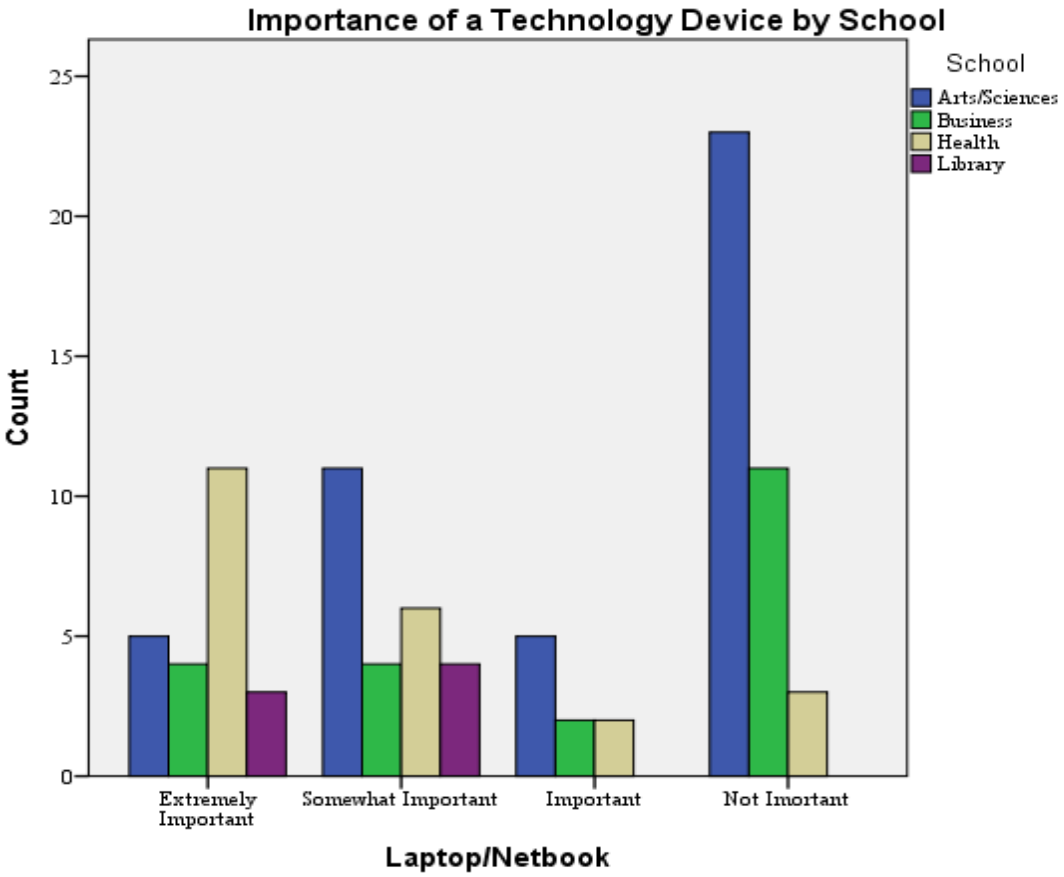
The issues explored in this question was the relationship between the school of affiliation and in faculty perceptions about the use of technology devices including the desktop computers, iPad/tablets, laptops, smartphones, or e-readers. Questions 1 and 7 were examined to analyze this relationship. Figure 4 displays the cross tabulations for each of the devices. There was a significant relationship at the .05 level between the school of affiliation and the importance of the desktop PC χ^2 (9) = 23.344, p = .005; the Phi coefficient was .509. The perceived importance of

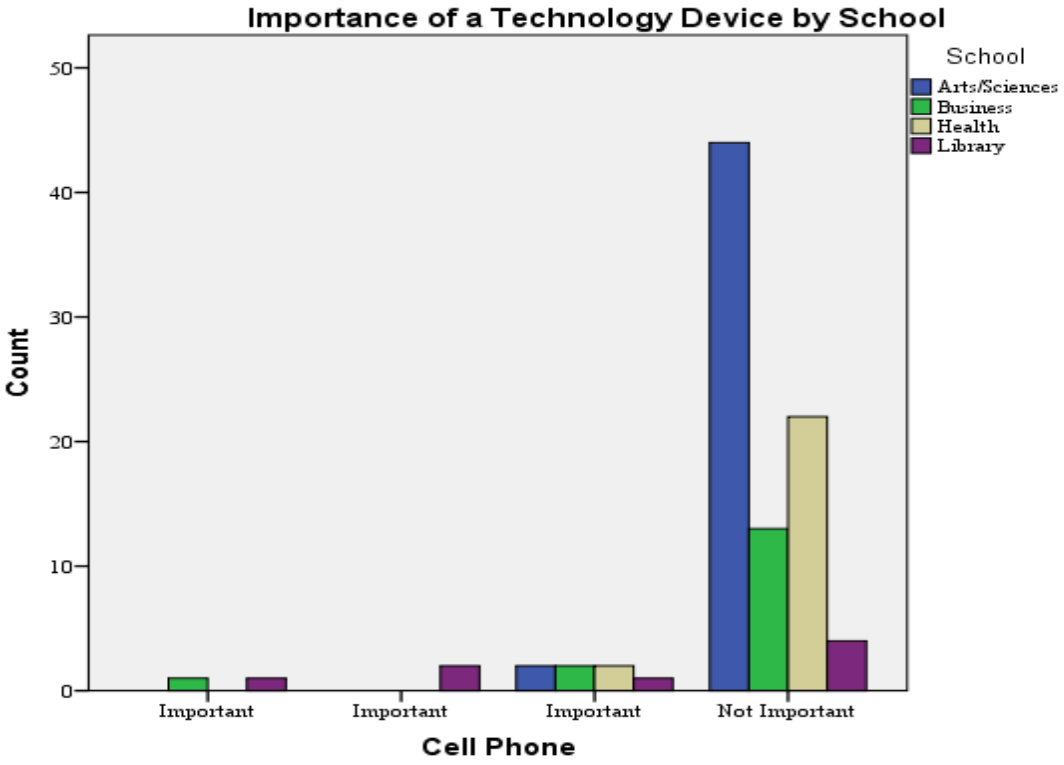
the iPad/tablet was also significant, χ^2 (9) = 34.285, p < .00; the Phi coefficient was .595. The importance of the laptop/netbook was significant, χ^2 (9) = 23.165, p = .006; the Phi coefficient was .496. The cell phone and its importance among the schools was significant, χ^2 (9) = 31.486, p <.0; the Phi coefficient was .579. The association between the e-reader's importance and the school was not significant χ^2 (9) = 14.795, p =.097; the Phi coefficient was .401.

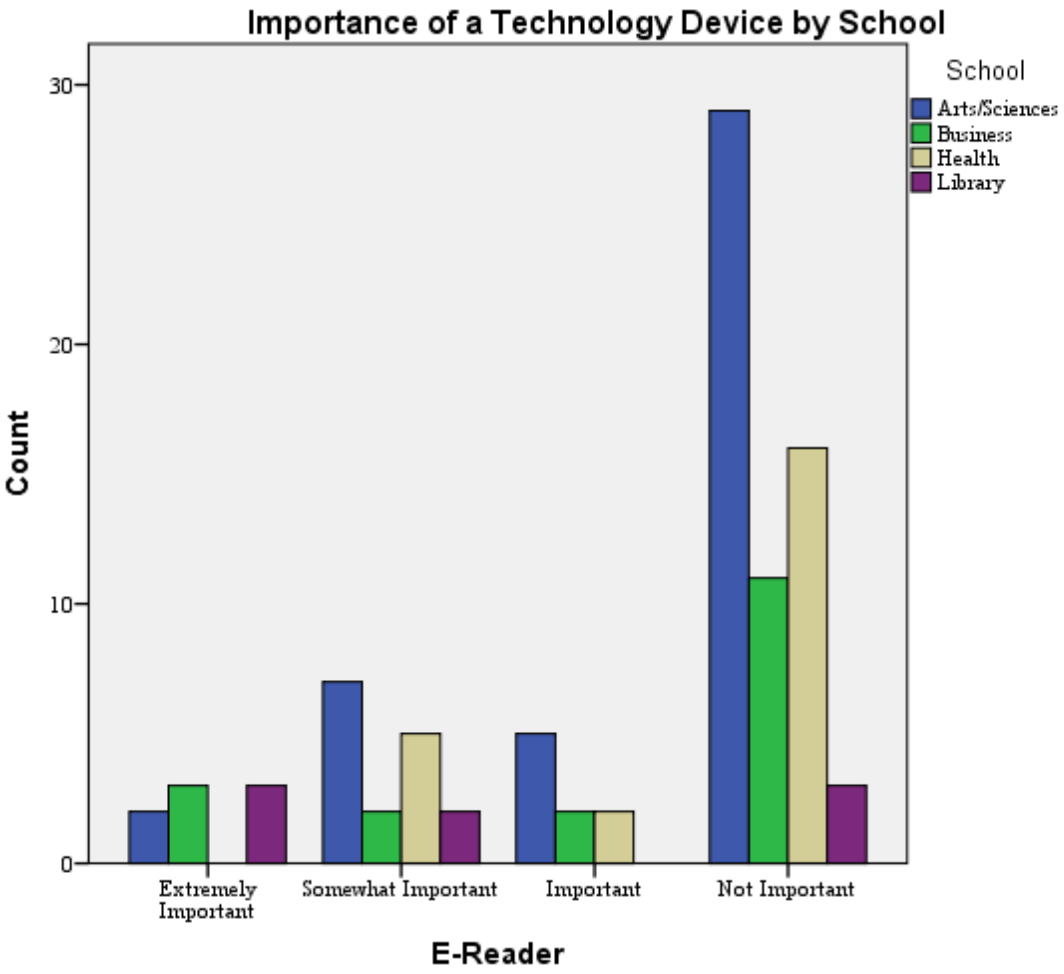
Figure 4




Importance of a Technology Device by School
School
Arts/Sciences
Business
Health
Library
Count
40
30
20
10
0
Extremely Important
Somewhat Important
Important
Not Important
iPad/Tablet


Importance of a Technology Device by School
School
Arts/Sciences
Business
Health
Library
Count
25
20
15
10
5
0
Extremely Important
Somewhat Important
Important
Not Imortant
Laptop/Netbook

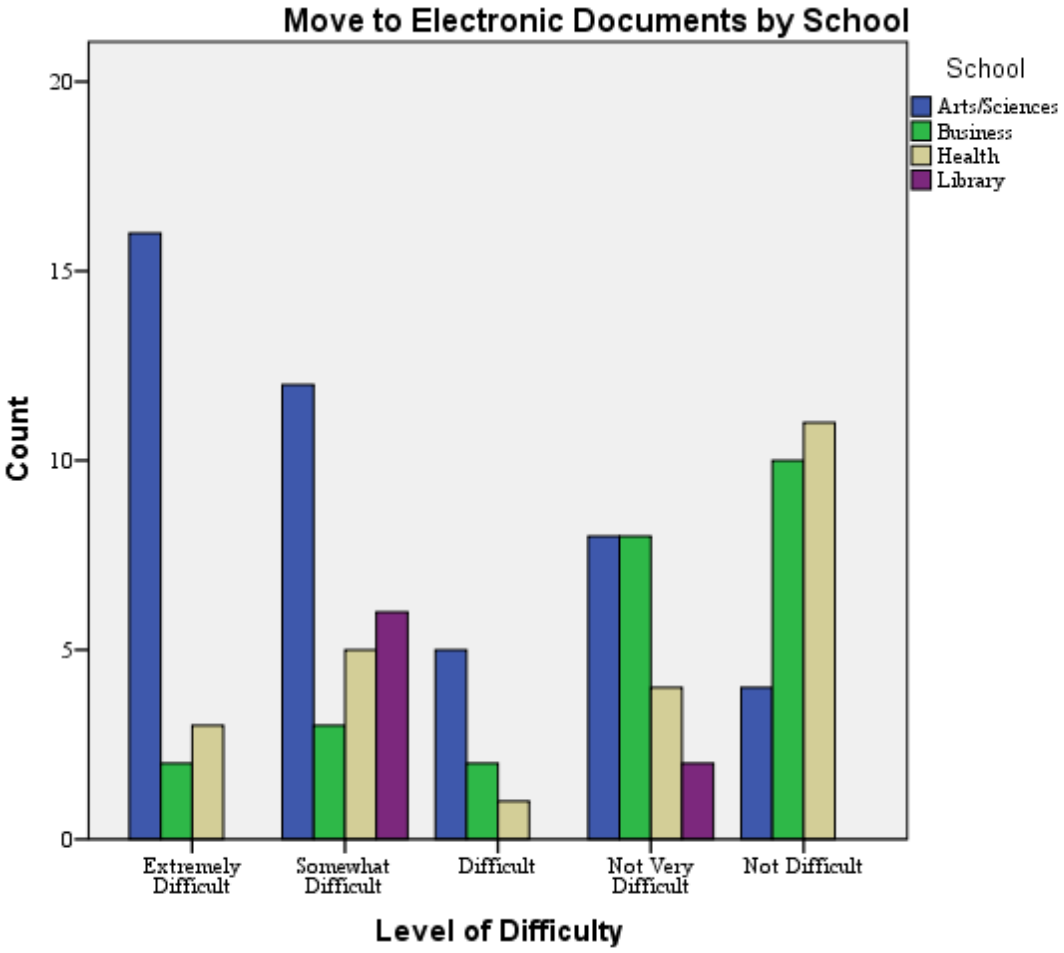

Importance of a Technology Device by School
School
Arts/Sciences
Business
Health
Library
Count
50
40
30
20
10
0
Important
Important
Important
Not Important
Cell Phone



Research Question 5

This question examines the relationship between the school of affiliation reflected and in the view of the difficulty in moving toward electronic documents instead of printed handouts. Questions 1 and 11 were used for this part of the study. Figure 5 shows the relationship between these variables. This association was significant at the .05 level χ^2 (12) = 35.804, p < .0; the Phi coefficient was .592.

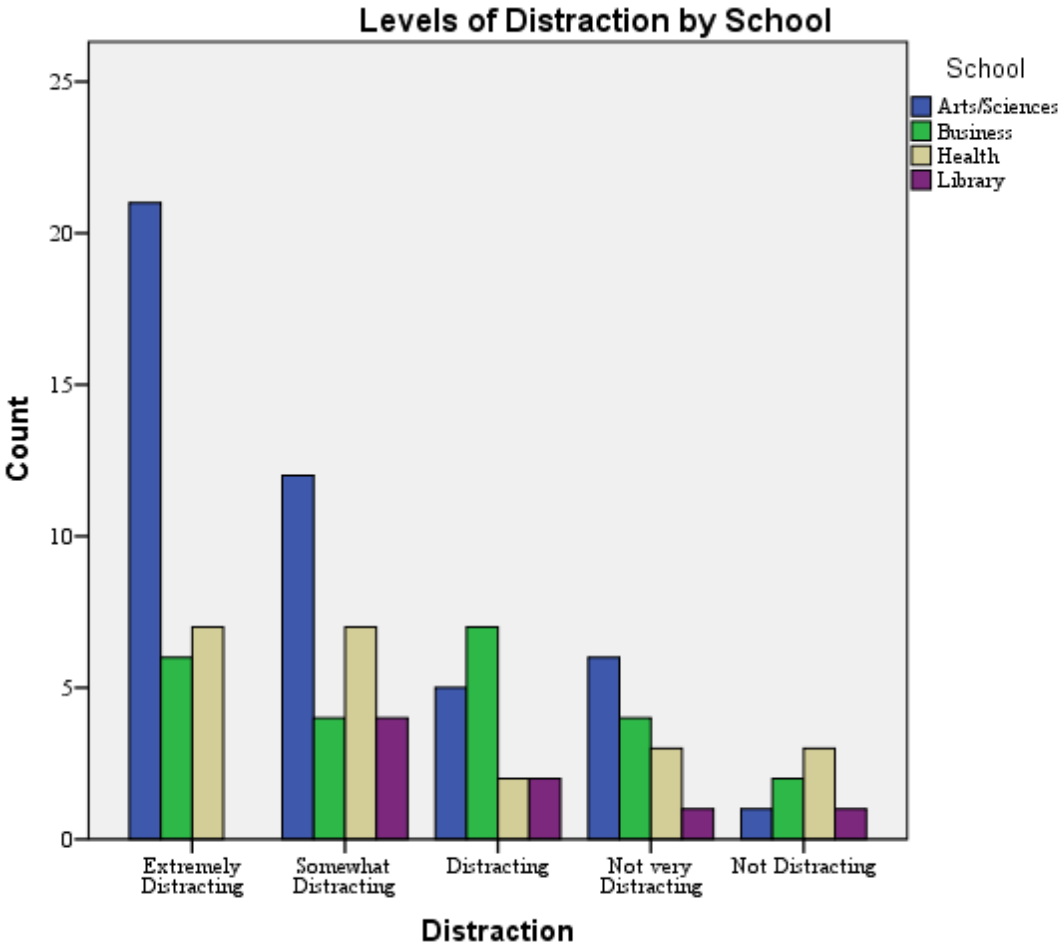
Figure 5



Research Question 6

This question investigated the relationship between faculty affiliations and whether technology devices are viewed as distractions. Questions 1 and 12 were used to examine this association. Figure 6 presents the bar graph of the data. This relationship was not significant at the .05 level χ^2 (12) = 15.72, p =204; the Phi coefficient was .401.

Figure 6

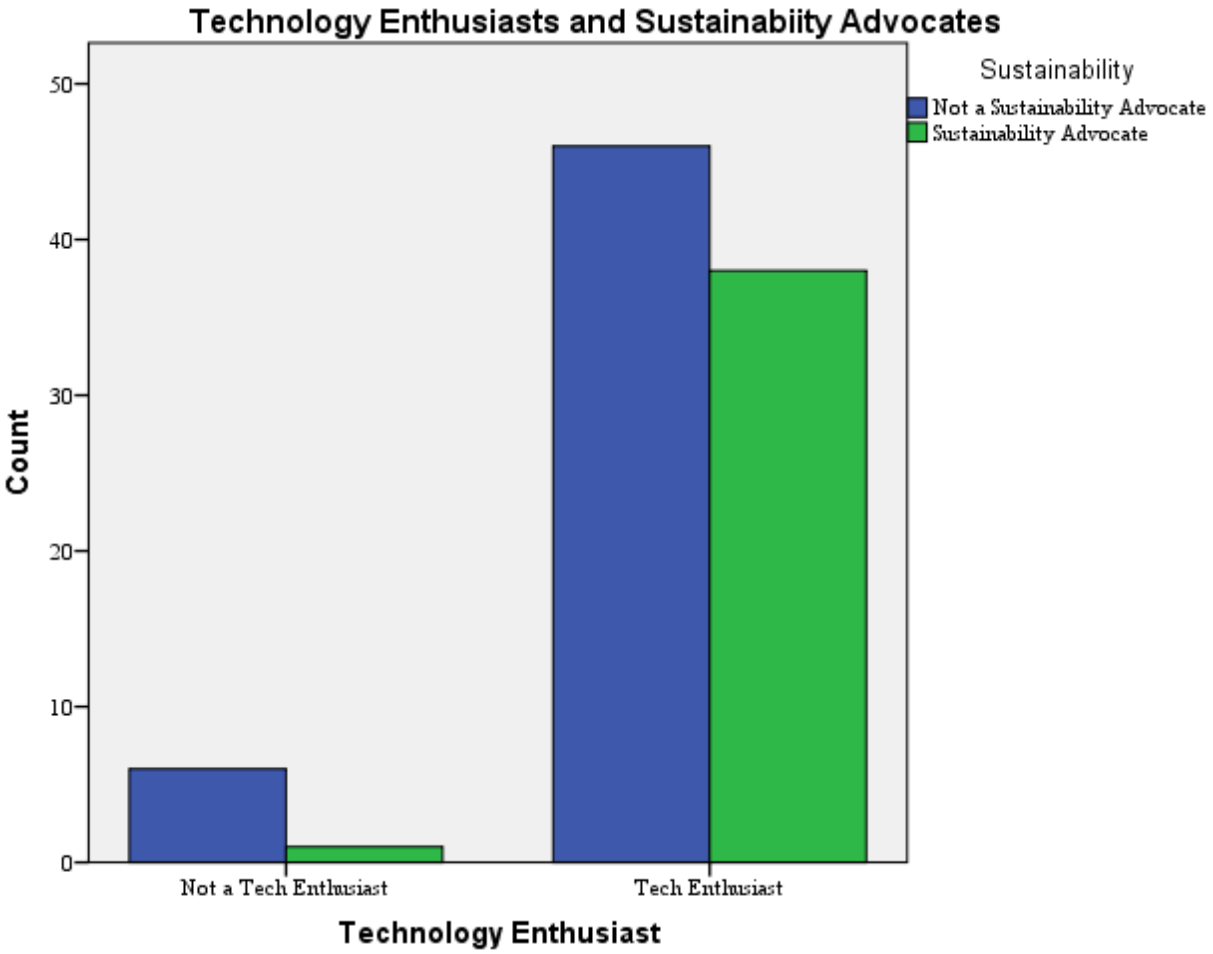


Research Question 7

This question explores the relationship between faculty who are technology enthusiasts and faculty who are sustainability advocates. Questions 4, 5, and 10 were applied to this analysis. If a faculty member answered Question 4 as extremely important, somewhat important of important, they were classified as a technology enthusiast. If a faculty member answered Question 5 as yes and Question 10 as do not require any printed handouts they were categorized as a sustainability advocate. Figure 7 displays the results of the comparison. The association

between technology and sustainability advocates was not significant at the .05 level χ^2 (1) = 2.528, p = .232; the Phi coefficient was .167.

Figure 7



Findings

The responses to the questionnaire revealed information that is useful in profiling the attitudes of faculty toward the use of technology in the classroom and in efforts toward sustainability. Among the respondents in this study, faculty from the School of Arts and Sciences were the most represented. The average teaching years of all respondents was over 10 years. The importance of using laptops/netbooks in the classroom was given the highest level of importance.

The majority (86%) of faculty surveyed permit the use of electronic devices in their classrooms. Over 73% of faculty view technology as extremely important in their teaching. According to respondents in the survey, the technology devices that are important in the classroom in order of their importance are desktop PCs, laptop/netbooks, iPad/tablets, e-readers and cell phones. About 71% felt that electronic devices were distracting to some extent. The number of faculty who did not require the use of printouts in the classroom was 40%. A slight majority (50.7%) of the people surveyed felt that there would be some degree of difficulty in moving toward the use of electronic documents.

The findings of this study with regard to the research questions are varied. No significant relationship was found between differences in the use of technology based on the faculty member's school affiliation. Faculty who teach in technology-based programs as could be expected in the School of Business or Library Services did not perceive the use of technology differently from faculty in other schools. There was a significant difference detected in faculty affiliations and the faculty member's view of the importance of technology to the learning process in the classroom. Differences were also found in the perception in the type of technology (desktop computers, iPad/tablets, laptops, cell phones and e-readers) and its use based on school affiliation. There was a significant association between school affiliations and the importance of moving toward the use of electronic documents. Faculty affiliations did not impact whether technology devices were viewed as distractions. Finally, it was found that sustainability enthusiasts did not have significant relationship with technology advocates. Sustainability enthusiasts may or may not be technology advocates.

Conclusion

This study examined the role of technology in the classrooms at a university through a questionnaire. The importance of technology to both the teaching and learning processes were investigated. Effort was also made to capture information about sustainability. Four academic schools were included in the survey and provided a basis to determine if the school affiliation was a determining factor in the use, view, or role of technology in the classroom. Differences were found in the perception of the use of technology, in the use of technology devices, and in the difficulty in moving toward the use of electronic documents based on school affiliation. The results of this study were not able to conclude that sustainability enthusiasts would or would not also be technology advocates. It is recommended that addition research be conducted to examine the relationship between sustainability and technology.

References

Association for the Advancement of Sustainability in Higher Education. (2011). About the Association for the Advancement of Sustainability in Higher Education (AASHE). Retrieved December 1, 2011, from http://www.aashe.org/about

Edutopia. (2011). Why integrate technology into the curriculum? Retrieved December 1, 2011, from http://www.edutopia.org/technology-integration-introduction

Howell, D. C. (2002). *Statistical methods for psychology* (5th ed.). Pacific Grove, CA: Duxbury/Thomson Learning.

Klopfer, E., Osterweil, S., Groff, J., & Haas, J. (2009). The educational arcade. Retrieved December 1, 2011, from http://education.mit.edu/papers/GamesSimsSocNets_EdArcade.pdf

Simon, S. (2005). What is a phi coefficient? Retrieved November 2, 2010, from http://www.childrens-mercy.org/stats/definitions/phi.htm

Tamim, R. M., Bernard, R. M., Borokhovski, E., Abrami, P. C., & Schmid, R. F. (2011). What forty years of research says about the impact of technology on learning: A second-order meta-analysis and validation study. *Review of Educational Research, 81*(1), 4-28.

United States Environmental Protection Agency. (n.d.). Sustainability basic information. Retrieved December 1, 2011, from http://www.epa.gov/sustainability/basicinfo.htm

APPENDIX A

Survey Questions

1. Which school are you working at?
    a. School of Arts and Science
    b. School of Business Administration
    c. School of Education and Human Services
    d. Malek School of Health Professions
    e. Library and Learning Services
2. What is your faculty status at the Marymount University?
    a. Professor
    b. Associate Professor
    c. Assistance Professor
    d. Term Appointment
    e. Adjunct or Instructor
3. How many years have you been teaching college students?
    a. More than 10 years
    b. 5 to 10 years
    c. 2 to 5 years
    d. Less than 2 years

4. How important do you think classroom electronic equipment, such as desktop PC, projector, or SmartBoard, is to your teaching in the classroom?
    a. Extremely Important
    b. Somewhat Important
    c. Important
    d. Not Very Important
    e. Not Important
5. Do you allow laptops or other electronic devices in your classroom for students to take notes and access files electronically? Y N
6. How important do you think laptops or other electronic devices are to the learning process in the classroom?
    a. Extremely Important
    b. Somewhat Important
    c. Important
    d. Not Very Important
    e. Not Important

For Questions 7 through 11 use the scale below to rate each of the following devices listed in the table. How would you rate the importance of each of them in the classroom? (1-Extremely Important, 2-Somewhat Important, 3-Important, 4-Not Important)

TABLE 1: Electronic Devices

Questions	Extremely Important	Somewhat Important	Important	Not Important
7. Desktop Computer	1	2	3	4
8. iPad/Tablet	1	2	3	4
9. Laptop/Netbook	1	2	3	4
10. Cell Phone	1	2	3	4
11. E-Reader	1	2	3	4

12. How distracting are those electronic devices when students use them in the classroom?

a. Extremely Distracting

b. Somewhat Distracting

c. Distracting

d. Not Very Distracting

e. Not Distracting

13. Who should have the authority of making the decision to allow or ban those electronic devices in the classroom?

a. Instructor

b. School

c. University

14. How frequently do you require students to bring printed handouts to class?

a. Once a Week

b. A few times a Month

c. A few times a Semester

d. Don't require at any time

15. How difficult would it be to make the shift to using electronic documents/information instead of printed handouts in the classroom?

a. Extremely Difficult

b. Somewhat Difficult

c. Difficult

d. Not Very Difficult

e. Not Difficult

Analyzing Retention from Participation in Co-curricular Activities

Roger Blake
University of Massachusetts Boston
Management Science and Information Systems
100 Morrissey Blvd
Boston, MA 02125-3393
617-287-7692
roger.blake@umb.edu

Paul Mangiameli
College of Business Administration
University of Rhode Island
Kingston, RI 02881-0802
401-884-4217
mangia@uri.edu

ABSTRACT

The impact of student retention is widely recognized by colleges and universities as extending well beyond financial, and retention is an issue that researchers have increasingly been turning to. Recently many have been focused on analytic methods, especially predictive models based on data mining techniques. Generally, these models use data on demographics, academic performance, socio-economic backgrounds, standardized tests, or financial aid. Student engagement is known to influence retention, but also can be more difficult to measure and is used less often. When student engagement is used, it is generally represented by the extent to which student support services such as advising are utilized. We explore a new approach to analyzing the influence of student engagement on retention and demonstrate how data mining models for retention might be developed based on data from participation in a program of co-curricular activities named MAP. Our research uses cluster analysis to show there are three distinctly different patterns of participation, and that participants tend to be in one of three groups, which we termed "High achievers", "Meets expectations", and "Retention risks." We find statistically significant differences between each of these patterns and how those individuals most likely to be at risk of attrition might be identified.

INTRODUCTION

Student retention has an enormous impact on colleges and universities that goes well beyond the financial. Retention and student attrition affect the reputation, rankings, culture, enrollment decisions, and the viability of educational institutions. Even with the recognition of how importance retention is and considerable efforts to address it, the percentage of students who enroll and complete their undergraduate degrees has changed little [9], and nationwide the percentage of students who complete a four year degree is only 35.3% for public four year

institutions and 40.7% for those that are private [1]. It is then not a surprise that researchers have increasingly turned their attention to this topic. The methodologies most used in these studies are surveys and data mining; of these, data mining has recently been gaining prevalence.

Prior research of predictive data mining models for retention has generally used a combination of data on academic performance, demographics, test scores, socio-economic background, and family education levels. While student engagement is known to influence retention, when it has been used in data mining models, it has generally been through measures of the utilization of student support services such as tutoring or academic advising. Our research explores a new approach to evaluating how student engagement influences by incorporating participation in a program of co-curricular activities.

This paper explores how participation in that program, named the Management Achievement Program, or MAP, might be used to categorize and predict student retention. The purpose of MAP is a program is to build professional skills and student engagement, and that end it has been successful. MAP is required for all students, and because of that, the opportunity to test whether MAP increases the sense of involvement or engagement is limited – there is no control group for such a test. However, as part of administering MAP an array of data regarding participation and events is gathered. This research investigates how that data might be utilized for data mining models to predict retention, and models potentially capable of making that determination very soon after enrollment when attrition rates are known to be highest and accurate predictions the most valuable.

BACKGROUND

Prior work

Data mining models used to predict retention from prior research are most often built with data for a combination of demographics, financial aid, family educational backgrounds, performance measures, standardized tests, and faculty experience. For example, Nandeshwar, Menzies, and Nelson [8] note there is much room for improving predictive models of retention and provide a comprehensive review of prior research. They hypothesize three factors as having significant influences on retention: financial aid, academic performance, and faculty tenure and experience, finding socio-economic and academic performance to be most significant. In another example, starting with demographic data such as age, performance measures such as SAT scores, chosen majors, work study history, financial aid data, and data on scholarships, Defen [3] focused on ensemble methods to evaluate a suite of data mining algorithms. In concluding which ensemble methods are better, Defen commented that the results from the models might have been improved if more data were available, particularly for social interaction. The participation data used for our research may be useful in this regard because a number of MAP events involve student clubs and volunteer activities, although we hold that investigation for further research.

There is considerable evidence that student engagement influences outcomes including critical thinking, cognitive development, self-esteem, satisfaction, and the factor that is most central to our research, retention [13]. When student engagement has been incorporated in data mining models of retention, engagement has often been represented by the extent to which student

support services have been utilized. For example, several models were compared using data on support service utilization and publicly available data on demographics, family education levels, course enrollments, and financial aid. That study ultimately concluded that utilizing student support services and successful completion of foundation courses were the most significant influences on retention [4].

Another example of measuring student engagement through the utilization of services, and a precedent for using cluster analysis to analyze its patterns as we do, is from a typology based on clustering data on utilizing student services such as academic advising, skill labs, and student organizations, combined with demographic data for 320,000 students [11]. With a much wider scope and population for analysis, the typology from their cluster analysis consisted of 15 clusters rather that the three found from ours. However, like our research, distinct clusters of participation could be identified, and like our research, the significance of these patterns for analyzing and improvement retention was a motivation for the study.

MAP

This research draws on a professional development program called the Management Achievement Program (MAP). MAP was not developed for the purpose of gathering data for analyzing retention. Instead, it is a program designed to develop and enhance the professional demeanor of its participants. The rationale for introducing MAP was that professionalism can best be developed and demonstrated when participants engage in co-curricular activities, in contrast to the assimilation of content delivered only through lectures or seminars. The philosophy of MAP is based on Experiential Learning and its premise that learning is "a cycle driven by the resolution of the dual dialectics action/reflection and experience/abstracts" [5], a philosophy embedded in how the program is defined:

> "An engaging and comprehensive program designed to develop and enhance each student's professional demeanor, build competencies for academic success, increase involvement in the College and local business communities, and allow the opportunity for students to personally synthesize their academic and professional goals and experiences."

MAP is a requirement for all undergraduate students in business administration at a state university. Students in MAP are required to select and participate in events and activities designed to build professional and career skills. Among the types of events and activities regularly offered are career workshops, seminars, forums, company visits, presentations by senior executives, student clubs, and service learning activities such as volunteer work.

Upon admission into the college, each participant's record is assessed to determine the number of MAP miles to be required for completing the program. Miles are used as the metaphor for MAP participation, and once the MAP miles requirement has been established, a participant can start to earn miles to fulfill that requirement by participating in MAP events. Each event offers a number of miles depending on the involvement and initiative it requires. For the majority of events participants can earn 50 miles; an example of such an event would be a workshop on interviewing. Events such as competing in a college-wide case competition can earn as many as 200 miles. Some events also require participants to record their reflections to summarize what

they learned and how they envision applying it to their academic or professional careers. These written responses are also a source of valuable information that might be used in conjunction with the data in this study, but beyond the scope of this paper and part of further work.

The MAP program operates as follows. After matriculation and prior to their first semester, each participant is given a bar-coded identification card and access to a web-based portal for reviewing upcoming events, registering for specific events, entering reflections on events attended, and checking the status of MAP miles and accounts. Because MAP events are independent of course schedules and often held outside of classrooms or off-site, attendance and participation is recorded through a bar-coding scanning system which is regularly uploaded to a back-end system.

Figure 1 shows the infrastructure for supporting MAP; it should be noted that this figure depicts the operational component of MAP only.

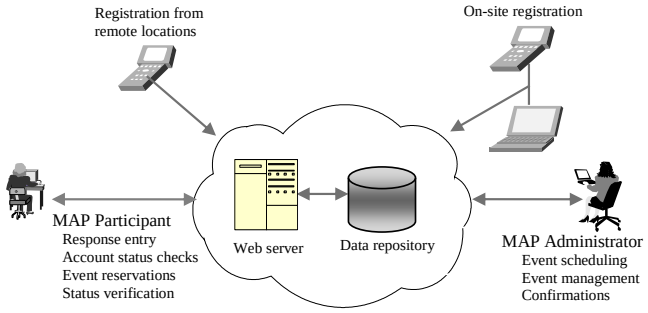


Figure 1. Management Achievement Program (MAP) operational infrastructure

MAP was implemented in the fall of 2006, and during that initial semester, 40 events were offered and attended by 202 participations. To maximize the program's benefits and offer participants flexibility, the number and type of events held each semester has grown significantly. During spring 2011, 2,466 participants were able to choose among 123 different events. MAP is now an integral part of the college's curriculum and well-accepted in its culture; from inception through the end of August 2011, there have been 1,020 MAP events attended by a total of 17,317 participants.

Integrating professional and soft skills into a curriculum is held to be desirable but difficult to implement [10]. MAP has brought these directly into the college's curriculum. Many institutions identify professionalism as a specific learning objective, but this is a learning outcome that is difficult to assess [7]. MAP has enabled professionalism and soft skills to be assessed with the use of methods such as Latent Semantic Analysis [2].

MAP has been successful as a program of co-curricular activities. But can MAP help predict student retention by analyzing how students participate in this program? That is purpose of this research, and the objective of the methodology discussed in the next section.

METHODOLOGY

Students are enrolled in MAP and made aware of their requirement for MAP miles and the operation of the program prior to beginning of their first semester with the college. The number of MAP miles they will need to earn to complete the program is based on a sliding scale, and is usually 250, 500, 750, or 1,000 miles. As noted earlier, students are then free to choose from the schedule of MAP events offered during and between semesters to complete their requirements, and the number of miles to be earned varies depending on the specific MAP event.

To evaluate whether or not participation in MAP might be related to retention and ultimately a successful academic career our first step was to determine if there exist groups whose progress through the program is similar, meaning to find whether distinct patterns of participation in MAP exist.

To find if there are any patterns in participation we conducted a cluster analysis. This analysis clustered participants using the percentage of their MAP miles requirement that was completed during each semester of their enrollment, starting from their first semester and extending through two years.

We ran numerous *k*-means cluster analyses and found that for values of *k* higher than 3, the analysis generally produced one or more clusters having a small number of observations that were not viable. After examining these runs, we determined that the most meaningful results were produced by a value of 3 for *k.* Using that setting, an individual cluster analysis was conducted for participants who starting in the same semester and held the same requirement for MAP miles. For example, there were four cluster analyses conducted for students starting in spring 2008, one each for those with requirements of 250, 500, 750, and 1,000 MAP miles.

A sample of the results from the cluster analysis is shown in Figure 2, which shows the average percentage of miles completed for participants who started in fall 2006 with 500 miles as their requirement for miles.

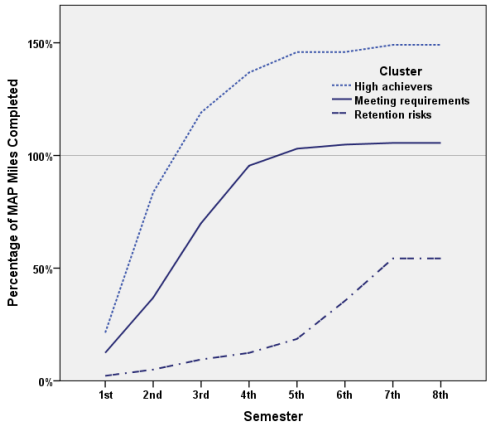


Figure 2: Clusters of MAP participation rates

The patterns in participation rates shown in Figure 2 can be interpreted as follows. Each line represents a cluster of participants, each showing the average percentage of miles participants

accumulated towards their requirements during each semester. For example, participants represented by the solid line in the middle reached approximately 15% of their requirement during the first semester, 30% during their second, 60% during their third, and completed their requirement by reaching 100% during their fifth semester.

The patterns of participation seen in Figure 2 are similar to those found for participants starting in different semesters and having different requirements for MAP miles. What can generally be seen in the cluster analyses is that in one cluster participants are going well beyond the required miles, we tentatively named that cluster "High achievers." The pattern for participations in another cluster seemed to be to fulfill just the essential requirements, and we termed that cluster "Meeting requirements." The pattern for the third cluster was distinctly different from the other two, and a pattern in which participants fell well short of meeting their required miles. This cluster we termed "Retention risks."

The apparently clear patterns that emerged from the cluster analysis, and the similarity of those patterns regardless of the semester a participant began their program or their requirements for MAP miles, led us to question whether the differences in participation rates for those clusters was statistically significant. To evaluate this we conducted one-way ANOVAs to test for significant differences between the participation rates in each cluster, i.e. between the percentages of miles completed over time for the participants in the three clusters.

The ANOVAs showed significant differences in participation rates; for example, for the 117 students starting in fall 2006 with a requirement of 500 miles, $p < .05$ as shown in Table 1.

	Sum of Squares	**df**	**Mean Square**	**F**	***p***
Between Clusters	.701	2	.351	12.056	.000
Within Clusters	3.315	114	.029		
Total	4.016	116			

Table 1: Test of mean participation rate differences, participants with a MAP miles requirement of 500 and starting in fall 2006

The three clusters visually appear to represent distinctly different patterns of participation in plots, but to test whether those are statistically significant differences we preceded with planned post-hoc tests after significant differences were indicated by the omnibus ANOVAs. Table 2 presents the results of the post-hoc test for the same participants as analyzed for Table 1, students who started in fall 2006 with a requirement of 500 MAP miles. It confirms the impression of distinct differences seen in plots, finding significant differences between each of the three clusters of MAP participation, $p < .05$.

		Subset for alpha = 0.05		
Cluster	**N**	**1**	**2**	**3**
3 ("Retention risks")	40	.0238		
2 ("Meeting requirements")	60		.1242	
1 ("High achievers")	17			.2618

Table 2: Results of Tukey-B test for homogeneous subsets of clusters

DISCUSSION

Of the three clusters from the results, the cluster "Meeting requirements" had the largest number of occurrences with slightly more than half of all participants. In general, these participants completed their miles requirements within four or five semesters. While we could not have expected this as a cluster, it is a logical result and consistent with participants generally completing the program. A finding of interest is that rather than simply fulfilling the requirement and ceasing participation, many participants in this cluster continued attending events and accumulating miles beyond the minimum required.

Participants in the "High achievers" cluster did not just meet their requirements for the program, but went far beyond them. During their eighth semester, on average they had exceeded requirements by an average of almost 50%. That this group would emerge as a cluster was not surprising, but the magnitude of the margin above the requirements and the size of the group - approximately 15% - were.

Some one-third of participants were in the third cluster which was termed "Retention risks." These are students who, unlike those in the other two clusters, have not completed their requirements for the MAP program after two years, and with an average of half of their required miles completed, do not appear close to be completing them. While some students in this cluster may have completed more than half of their requirements, and some may ultimately complete the program, most do not. There are many different paths that students can take which can make categorization for retention difficult; some may switch from full-time to part-time, withdraw mid-semester from their courses, or stop taking courses for a semester or two only to return later. Regardless, the participants in this cluster clearly stand out as potential concerns for retention.

Finally, the analysis and the separation of participation rates evident in Figure 1 during the first semester suggest that whether a student will ultimately be in either the "High achievers", "Meets requirements", or "Retention risks" cluster may be apparent very early in an academic career, even during their first semester. Retention during the first semester is especially critical, yet many studies rely on factors only available after it has ended, such as GPA. While it is outside the scope of this paper, we are currently investigating this as further research among other directions as discussed next.

SUMMARY

The importance of student retention is widely recognized by colleges and universities and the impact of student attrition widely felt. Researchers have increasingly turned to analytical methods and data mining in particular to explain and predict student retention. The factors impacting retention that these studies have explored include demographics, family backgrounds such as the educational levels of family members, performance indicators such as GPA and scores on standardized tests, and amount of financial aid have all be used as factors to analyze retention.

Our research investigates how predictive data mining models for retention can be developed based on a program of co-curricular activities named MAP. Prior research of data mining models generally uses factors such demographics, education levels of family members, performance

measures, amount of financial aid to predict retention. When student engagement or involvement is studied as a factor, it is most often represented by the degree to which student services such as academic advising or tutoring are utilized. To our knowledge, this is the first study to examine how an institutional program of co-curricular activities might be used to measure and predict student retention. This program was designed for the purpose of developing professionalism and soft skills, which can be difficult to integrate explicitly into curricula and very difficult to assess. Our research shows that there are distinctly different patterns in participation rates, and using cluster analysis finds that participants tend to be in one of three groups, which we termed "High achievers", "Meets expectations", and "Retention risks." Our research demonstrates that there is a statistically significant relationship between these participation patterns and retention.

This research is currently being extended in several significant directions. One of these is to build on the analytical methods presented in this paper. More specifically, we are augmenting the cluster analysis with other unsupervised techniques, such as association rules in a method from a seemingly unrelated but potentially applicable method developed in a study of ant colonization [6]. Other potential factors that might influence retention, including measures of performance such as GPA, are also being explored using supervised techniques to predict retention directly. Text analytics are the third category of techniques under development, which are being used to analyze the text of responses written by participants and the results integrated with those from the other techniques.

The statistically significant relationships found in our research raise the possibility that from participation in MAP, retention could be predicted very early in student's career. This would have significant implications because attrition rates tend to be highest during the first year and identifying students at highest risk for early intervention is thought to be one of the most effective means to prevent attrition.

Finally, we are currently pursuing the relationship our research to existing theories. Tinto [12] explores several theories of student engagement that can be tested by our research, and theories from other fields are also very relevant. In particular, customer retention is equally as important to businesses and student retention is to universities, and businesses have found methods of engaging customers to be highly productive. The problem of retention and customer "churn" have many similarities [3], and some of the existing theories used to explain and predict customer retention are theories that might be tested or adapted to see how they apply to retention models based on student participation as presented in this research.

REFERENCES

[1] ACT. What Works in Student Retention? . Fourth National Survey. Iowa City, IA2010.

[2] Blake R, Gutierrez O. A semantic analysis approach for assessing professionalism using free-form text entered online. Computers in Human Behavior. 2011;27:2249-62.

[3] Delen D. A comparative analysis of machine learning techniques for student retention management. Decision Support Systems. 2010;49:498-506.

[4] Fike DS, Fike R. Predictors of first-year student retention in the community college. Community College Review. 2008;36:68-88.

[5] Kolb AY, Kolb D. Experiential learning theory. In: Armstrong S, Fukami CV, editors. The SAGE Handbook of Management Learning, Education, and Development. London: SAGE Publications, Ltd.; 2009.

[6] Kuo RJ, Lin SY, Shih CW. Mining association rules through integration of clustering analysis and ant colony system for health insurance database in Taiwan. Expert Systems With Applications. 2007;33:794-808.

[7] Mayotte S. Online assessment of problem solving skills. Computers in Human Behavior. 2010;26:1253-8.

[8] Nandeshwar A, Menzies T, Nelson A. Learning patterns of university student retention. Expert Systems With Applications. 2011;38:14984-96.

[9] NCES. College persistence on the rise?: Changes in 5-year degree completion and postsecondary persistence rates between 1994 and 2000. In: NCES, editor. Washington, DC2005.

[10] Pfeffer J, Fong CT. The end of business schools? Less success than meets the eye. Academy of Management Learning & Education. 2002;1:78-95.

[11] Saenz VB, Hatch D, Bukoski BE, Kim S, Lee K, Valdez P. Community College Student Engagement Patterns: A Typology Revealed Through Exploratory Cluster Analysis. Community College Review. 2011;39:235-67.

[12] Tinto V. Research and practice of student retention: what next? Journal of College Student Retention: Research, Theory and Practice. 2006;8:1-19.

[13] Trowler V, Trowler P. Student engagement evidence summary. Lancaster, U.K.: The Higher Education Academy; 2010.

THE EFFECT OF WEB USABILITY ON USER'S WEB EXPERIENCE

Molly Herring, Bryant University, mherring@bryant.edu
Janet Prichard, Bryant University, (401) 232-6269, prichard@bryant.edu

ABSTRACT

The ease with which a website visitor can find what they need is positively correlated with visitor satisfaction [4]. Web usability is a field that studies what factors affect the visitor's ability to navigate through a website. Although there are publications outlining specific usability guidelines, many of them have little to no academic research to support the claim. The International Organization of Standardization and the U.S. Department of Health and Human Services (HHS) are two contributing organizations. HHS developed a list of 209 guidelines and rated each according to their strength of evidence (research-based supported, 5 – high, 1 – low). Using heuristic evaluation and usability testing, this study will provide additional research-based knowledge for those guidelines rated with a low strength of evidence. The study will show that research supports the claim that these guidelines positively impact web usability by aiding the user in finding what they need on a site and essentially increasing their satisfaction.

INTRODUCTION

As the internet continues to expand exponentially, there are those who believe specific standards are necessary in order to provide valid and valuable information in a guest-friendly manner. This is the field of web usability, whose main goal is to design and develop websites easy to understand and intuitively navigable. Web usability is an entirely outward facing, user-driven experience and forces companies and website owners to completely understand the goals, needs, and mindset of their visitors. Companies are beginning to understand that to ensure customer loyalty and return, they must cater their sites to their user's needs, as well as making sure they portray all the necessary information. As with many disciplines, the company must understand that their target audience does not have the same understanding of their services and often need to be convinced during the first impression – often within the first 10 second of visiting the page – that this website has the information they are seeking. Creating a bad first impression online can be more damaging than in person as it is more difficult to reach out to the customer; rather, they must accomplish the more difficult task of drawing the audience member to them.

Web usability is a relatively new field but has found stalwart support in numerous locations, including Jakob Nielsen, Ph.D., the International Standard of Organization (through the ISO 9241-151:2008 standard) and the Department of Health and Human Services (HHS). HHS conducted a study on what website features contributed to a visitor's ease of use when navigating through a website. These features became guidelines for the design of all sites within their Health and Human Services network. In addition to publishing the list of guidelines, they also rated each according to its relative importance (RI) and strength of evidence (SE). RI is scored 1-5 where 1 is of little important to usable web design and a 5 is of most importance. SE is

scored according to the amount of research performed in support of the guideline's inclusion on the list. HHS explicitly states that a low SE score is meant to encourage other researches to perform their own studies in order to contribute to the body of knowledge.

Contribution

This project will contribute to the web usability field by running a study on HHS guidelines that have a relative importance of 4 or 5 and strength of evidence rating of 1 or 2. This combination of scales will result in those guidelines rated highly important to web usability but have little or no academic research and evidence to support that claim. This will benefit the knowledgebase by providing further evidence for or against the inclusion of the guidelines in the HHS's publication. This work will be similar to other studies (ISO 9241-151:2008) in that it involves testing the usability of web elements. However, this work differs as each that each guideline used in the study has not yet been studied conclusively. There is little to no academic research performed on these guidelines. It is expected that the results will support the original decision to include these guidelines in the HHS publication.

LITERATURE REVIEW

"Web Development: A Visual-Spatial Approach" by Craig M. Baehr [1] discusses the importance of usability in the effective design and development of a website. Because the development process is iterative (a version is delivered, tested, feedback is given and changes are made) the testing process is repeated numerous times before the official release. Therefore, usability testing needs to occur not only once, but be integrated throughout the entire process. Baehr presents three different processes for testing that target different aspects of the site. Technical usability testing assesses the website's functionality across a variety of platforms and browsers. A usability worksheet can aid in organizing data and provide side-by-side comparisons of a site's output by different platforms. Visual-spatial usability testing focuses on how the user traverses the site, assuring that one can navigate through easily. Testing can be task-based – asking the user to perform a specific task and see how they do so – and evaluation-based where the tester examines the usefulness and flexibility of the tool. The third test assesses accessibility measures, ensuring that the site can reach the largest audience possible, include those with disabilities. For all of these measures, the same process can be followed: 1) establish goals, 2) develop testing instruments, 3) conduct the test, and 4) analyze the results.

The Health and Human Services guidelines are continually assessed and revised to "enable organizations to make more effective design decisions" [10, p. xvi]. Originally developed for the effective dissemination of information on Health and Human Services websites, the guidelines are useful for improving usability on all sites. There are currently 209 guidelines that, by the author's suggestion, can be used in numerous ways: a designer may follow the guidelines while creating a webpage, usability specialists can create a checklist when reviewing a site, or researchers can determine which guidelines require additional evidence and perform studies to acquire more. The guidelines have also been rated according to level of importance and strength of evidence. Level of importance describes which guidelines will aid in usability the most and should be first considered while designing a website. Strength of evidence relates to the amount of research performed about a particular guideline. Both research and expert opinion was

considered during the development of the guidelines, but low strength of evidence is meant to encourage further research. Please see Appendix 1 for an explanation of the strength of evidence scale. HHS does not provide an explicit definition for relative importance beyond a "5" is the most important element to consider including and "1" is the least.

HHS first began researching usability elements in 2004, compiling various guidelines from numerous different sources, including informal web usability websites, peer-reviewed research, and in-house standards and studies. The list totaled 500 originally but was reduced to 398 after the team took a first pass and combined or removed contradictory guidelines. Relative importance was decided among a group of 16 external reviewers consisting of eight web designers and eight usability experts. Guidelines that were determined to have little or no impact on the success of a site were removed, shrinking the list to 287. Eight usability researchers, practitioners and authors, all published and experienced in the field, created their own set of criteria for determining the strength of evidence for each guideline. More guidelines were removed as designers located graphical examples to be used in the published version. Twenty web designers partook in a card-sorting exercise to place guidelines into chapters and name them appropriately. These categories were then tested for usability. The final 2004 version had 187 guidelines; updates and revisions are made upon evaluating new research. The 2006 version updated 21 guidelines and added 22 more and the relative importance was revised based on a survey of 36 web professionals. These were edited by three independent groups of web professionals before publication.

ISO 9241-151, "Ergonomics of human-system interaction – Guidance on World Wide Web User Interfaces" [5] became an international standard in 2008 and promotes making websites "accessible to the widest possible range of users" [5, p. vi]. While this includes those with disabilities, the standard does not specifically discuss or target factors unique to that audience. While the entirety of ISO-9241 discusses usability of various agents, Part 151 focuses on the presentation of web content and a user's ability to navigate through a website. The World Wide Web creates unique usability issues because public websites are viewable by a wide range of audiences resulting in a variety of user capabilities, knowledge, and goals. Furthermore, there are numerous ways to access the internet, including multiple browsers and mobile agents that render pages differently. These guidelines can be applied to the internet, intranets, and extranets, but should not be applied indiscriminately to mobile devices. Similar to the HHS guidelines, this standard's target audience includes web professionals, developers, content providers, usability experts and other users who are impacted by or interested in web usability. As such, the standard provides a sample procedure and checklist for assessing a website's usability.

Not only are various usability guidelines occasionally contradictory with each other, but they may contradict other web standards. Visser and Weideman explore this idea in "An empirical study on website usability elements and how they affect search engine optimization," targeting web elements that play a role in both usability and search engine optimization (SEO) [11]. Both concepts are important to consider when designing and developing a website, but they target two different audiences: the former process aids search engine spiders in locating and indexing websites while the latter instead targets a human audience, whose goals upon visiting the site are often many and varied. For example, a search engine spider does not consider the aesthetic value of the page, while the human user may be more attracted to those sites that contain professional

graphics and evocative imagery and colors. The authors note that research suggests adding a page for a site's privacy policy, about us, and testimonials to increase the credibility of the site and encourage the user's trust. However, these pages do not directly relate to the topic of the website and therefore are considered extraneous to SEO.

Visser and Weideman chose four website attributes that were a source of contradiction between usability and SEO: trust and credibility, single page view and content, keywords, and use of images [11]. A control site was created by a web designer with little knowledge of usability and SEO guidelines and an experimental site was developed according to usability attributes, without any consideration for SEO. The study measured the level of traffic each site received, how many pages each visitor viewed, and how many visitors actually purchased the product that was being sold. From the results, Visser and Weidemann concluded that "usability is a prerequisite for effective website design" [11, p. 1] because a higher percentage of visitors to the experimental site submitted a contact form ('conversions') than the control site. Further research would include a second experimental website that focuses on SEO elements and ignores usability guidelines.

METHODOLOGY

A preliminary list of all guidelines was identified accordingly: those selected were rated by HHS with a "1" or "2" in their strength of evidence (SE) scale and were also rated "4" or "5" in their relative importance (RI) scale. Six specific guidelines were chosen for testing in an effort to provide additional strength of evidence. Traditional usability testing methods work with a developed site or site currently in development and test whether or not the site meets a set of heuristics and needs of the users. This study will test the set of standards used to develop a usable website. Two travel industry websites, travelocity.com and expedia.com, were chosen for a heuristic evaluation, which will locate how the sites operate in accordance with the 6 guidelines chosen for testing. Usability testing with human subjects will then determine how much effect these guidelines play on the usability of the website and how important the user rates the standards.

Guidelines

There are 60 guidelines that were given a 1 or 2 rating on the SE scale (please see Appendix 2 for this list). Of these, five had a 5 RI rating. Twenty-one guidelines were rated a 4 and the remaining 34 scored 3 or below. Six guidelines have been chosen that can readily be tested using currently existing websites, without the need for developing an experimental website. Below are the six chosen guidelines, as well as a summary of HHS's current findings:

- *Develop Pages that Will Print Properly*: a website should be able to print onto the standard 8.5 x 11" piece of paper without cutting off any of the content.
- *Provide Feedback on User's Locations*: provide visual cues as to where the user currently is in the hierarchy of the website. This may include breadcrumb navigation or colored tab structure.

- *Provide Descriptive Page Titles*: search engines use page titles to index sites so using descriptive, clear titles will help the user determine if this is a useful website to visit.
- *Link to Related Content*: the related content may be useful to the user.
- *Label Pushbuttons Clearly*: a user should know exactly what will happen when they click on a pushbutton.
- *Provide a Search Option on Each Page*: while a site's search engine shouldn't be relied upon too heavily, a user should not always have to manually scrounge around the site to find what they need.

This list is also located in Appendix 3.

Heuristic Evaluation

In order to minimize bias in usability testing, two sites were chosen from an industry that is unlikely to be frequently visited by the average participant in the study. The websites `travelocity.com` and `expedia.com` are known travel industry websites used to locate and book flights, hotels, and other related travel necessities. The travel industry was chosen because while it is likely that many of the test subjects will have heard about or infrequently used these sites, they will not be experts.

A heuristic evaluation was then used judge the compliance of a site to the 6 guidelines in question. All the pushbuttons were located and then analyzed for their clarity. It was noted that both sites adequately located the pushbutton next to the form it was submitting and was labeled in a manner that clearly denoted its action. Both sites were similarly structured and executed and displayed flight searches in a comparable manner, with minor differences. Two sites were chosen to ensure that it is the guidelines being studied for importance and effect on usability, not the individual site's usability. In the future, it would also be useful to choose two sites that fall on opposite ends of the usability spectrum. If a user is able to more easily navigate through a site that follows the heuristics instead of one that does not, it will provide credence for the HHS's standards.

Usability Testing

The study will be conducted with undergraduate students from a small private college in the Northeast and will need to account for demographic and stratification factors. Usability testing involves requesting participants to perform specific tasks on the located websites. A worksheet outlining each of these tasks will be given to each participant. Some tasks will be purposefully vague, forcing the user to interact with the site and locate the necessary page or information themselves. The sessions will target the user's ability to complete the task, the length of time it took them to do so, and their level of frustration in performing the task. An exit interview will follow to acquire any information that could not be gathered by observing the user performing the task. Questions include their opinion of the task, their perception of how easy or difficult it was to complete, and if they believed inability or difficulty completing the task would force visitors away from the site.

The outline for the preliminary usability session is included in Appendix 4.

CONCLUSIONS

The Department of Health and Human Services provides a plethora of guidelines available for the further research. Reducing this list to a manageable handful for the scope of this study is a process itself and could, in the future, be performed differently to test and produce different results. Therefore, replicating this study, both to verify results and provide additional evidence for HHS, is highly encouraged.

REFERENCES

[1] Baehr, C. M. *Web Development: A Visual-Spatial Approach.* Upper Saddle River: Pearson Education, 2007.

[2] Bevan, N. Guidelines and Standards for Web Usability. *HCI International* (p. 10). London: Lawrence Erlbaum, 2005.

[3] Chisholm, W., Vanderheiden, G., & Jacobs, I. Web content accessibility guidelines 1.0. *Interactions*, 2001, 8(4).

[4] Institute for Dynamic Educational Advancement. "Finding Information: Factors that Improve Online Experiences", 2008.

[5] International Organization for Standardization. "ISO 9241-151: Ergonomics of Human-System - Guidance on World Wide Web User Interfaces", 2008. Retrieved September 23, 2011, from `http://www.ilet.yildiz.edu.tr/sources/etkOrtTas/readings/2008_ISO-9241-151-Ergonomics%20of%20human-system%20interaction%20Part%20151%20Guidance%20on%20World%20Wide%20Web%20user%20interfaces.pdf`

[6] Johnson, J. "Web Usability in 2008: Mediocre", 2008. Retrieved September 27, 2011, from UI Wizards, Inc.: `http://www.uiwizards.com/WebUsability2008.pdf`

[7] Nielsen, J. *Designing Web Usability.* Indianapolis: New Riders, 1999.

[8] Nielsen, J. *Eyetracking Web Usability.* Indianapolis: New Riders, 2009.

[9] Nielsen, J. *Prioritizing Web Usability.* Indianapolis: New Riders, 2006.

[10] U.S. Department of Health and Human Services. "Research-Based Web Design & Usability Guidelines", *Washington, D.C.: U.S. Government Printing Office*, 2006.

[11] Visser, E. B., & Weideman, M. An Empirical Study of Website Usablity Elements and How They Affect Search Engine Optimisation," *SA Journal of Information Management*, 2011, 13 (1).

APPENDIX 1 - HHS STRENGTH OF EVIDENCE SCALE [10]

5 – Strong Research Support ●●●●●
- Cumulative and compelling, supporting research-based evidence
- At least one formal, rigorous study with contextual validity
- No known conflicting research-based findings
- Expert opinion agrees with the research

4 – Moderate Research Support ●●●●○
- Cumulative research-based evidence
- There may or may not be conflicting research-based findings
- Expert opinion
  - Tends to agree with the research, and
  - A consensus seems to be building

3 – Weak Research Support ●●●○○
- Limited research-based evidence
- Conflicting research-based findings may exist
  - and/or -
- There is mixed agreement of expert opinions

2 – Strong Expert Opinion Support ●●○○○
- No research-based evidence
- Experts tend to agree, although there may not be a consensus
- Multiple supporting expert opinions in textbooks, style guides, etc.
- Generally accepted as a 'best practice' or reflects 'state of practice'

1 – Weak Expert Opinion Support ●○○○○
- No research-based evidence
- Limited or conflicting expert opinion

APPENDIX 2 - PRELIMINARY LIST OF GUIDELINES

#	Title	Relative Importance	Strength of Evidence
1:5	Set and State Goals	4	2
2:12	Develop Pages that Will Print Properly	4	2
3:1	Comply with Section 508	5	2
3:2	Design Forms for Users Using Assistive Technologies	5	2
3:4	Enable Users to Skip Repetitive Navigation Links	4	2
3:5	Provide Text Equivalents for Non-Text Elements	4	2
3:6	Test Plug-Ins and Applets for Accessibility	4	2
4:1	Design for Common Browsers	4	2
4:2	Account for Browser Differences	4	2
4:3	Design for Population Operating Systems	4	2
4:4	Design for User's Typical Connection Speed	4	2
5:2	Show All Major Options on the Homepages	5	2
7:1	Provide Navigational Options	4	2
7:4	Provide Feedback on User's Location	4	2
9:2	Provide Descriptive Page Titles	4	2
10:2	Link to Related Content	4	2
10:4	Avoid Misleading Cues to Click	4	2
10:7	Designate Used Links	4	2
11:2	Format Common Items Consistently	4	2
13:2	Label Pushbuttons Clearly	5	2
13:4	Do Not Make User-Entered Codes Case Sensitive	4	2
15:4	Define Acronyms and Abbreviations	4	2
15:5	Use Abbreviations Sparingly	4	2
16:3	Ensure that Necessary Information is Displayed	5	2
17:3	Make Upper- and Lowercase Search Terms Equivalent	4	2
17:4	Provide a Search Option on Each Page	4	2

APPENDIX 3 – CHOSEN GUIDELINES

#	Title	Relative Importance	Strength of Evidence
2:12	Develop Pages that Will Print Properly	4	2
7:4	Provide Feedback on User's Location	4	2
9:2	Provide Descriptive Page Titles	4	2
10:2	Link to Related Content	4	2
13:2	Label Pushbuttons Clearly	5	2
17:4	Provide a Search Option on Each Page	4	2

APPENDIX 4 - USABILITY SESSION OUTLINE

About You

- Demographics: Gender, Major, Age, College year, Home state,:
- How often do you use sites similar to Travelocity, Expedia, Priceline, etc?
- Have you ever participated in usability testing before?
- What browser do you commonly use? (If possible, get version as well)

Scenario

You are a recent college graduate who has accepted employment in Durham, NC. You need to fly home on September 14th to attend a sibling's wedding. You will be returning September 16th. Find a roundtrip flight from Durham, NC to Bar Harbor, ME for these dates. Find the cheapest flight.

Also, complete these tasks:

- Choose an alternative airport near Bar Harbor, ME
- Locate a hotel in ME
- Choose the cheapest flight and print the results
- Return to the homepage
- Locate the Return Policy

Questions During Testing

- What page are you currently on?
- What is the page title of the page you are on?

Questions After Testing

- How easy did you find the site to navigate? What, if anything, made navigation difficult?
- What is a pushbutton? Did you find the one's on this site to be labeled clearly? How important are clearly labeled pushbuttons?
- The material that printed, is it clear and readable?

Find content that isn't what you are currently looking for. Is it related to your search? Do you find this helpful or annoying?

Institutional Incentives on IT Usage: The Principal-Agent Problem

Alexander Pelaez, Information Technology & Quantitative Methods
Hofstra University

Introduction

Information Technology has had a profound impact in changing many industries from adjusting firm size and decision rights (Gurbaxani & Wang, 1993) to enabling organizational change and increasing productivity (Dedrick, et. al 2003). For technology to effect change, users and firms must implement technology, adopt it and use it effectively. Research has demonstrated particular antecedents of adoption and intention to use technology such as usefulness and ease of use (Davis, 1989). Simply looking at adoption of technology potentially may not reveal an accurate prediction of actual usage. Therefore, more recently, researchers have focused their attention on antecedents of actual usage and its importance in effecting change. Unfortunately, this research had failed to deliver a consensus on the actual definition (Burton-Jones & Straub, 2006). Therefore, searching for the effect of Information Technology on any industry requires an understanding of its impact with respect to usage and therefore a clearly defined usage construct must be created that is meaningful.

Even if firms within an industry can benefit from Information Technology, there exists an amount of uncertainty and risk with respect to the investment in technology. Smaller firms within an industry could reap the benefits of technologies that are implemented in larger firms, however, due to the size these investments, they are sometimes unable or unwilling to invest capital and time (Tassey, 1996). Economic researchers have shown that the increased uncertainty, or risk, is negatively correlated with investment and that firms with greater risk aversion will tend to have lower outputs and inputs (Ghosal & Loungani, 2000). It could thus be inferred that the level of investment in IT may be a driving force for the firm; however, the driving force is actual usage of the IT artifact and not necessarily the investment in IT (Devaraj & Kohli, 2003). Therefore, other factors, beyond investment are at work toward actual usage. Unfortunately, firms that are risk averse will not make investments in technology and could be at a disadvantage.

Ultimately, for certain industries to grow and remain competitive, especially in a global environment, it is exactly this uncertainty that government or institutional incentives are designed to help alleviate (Tassey, 1996). Government intervention generally occurs when market failures threaten the public welfare (Taylor, et. al 2005). Governments have numerous tools at its' disposal to respond to market failures (Tassey, 1996) or protect the welfare of the people. It is the decision of what tool to use given a certain set of circumstances that should dictate the government's response coupled with its desired effect (Tassey, 1996). The challenge is the understanding the type of market failure that is occurring and matching the appropriate response (Tassey, 1996).

Government or institutional intervention is specifically designed for a need; however, governments never incentivize for a single firm. Intervention, with respect to technology adoption, is meant for an industry, and therefore, the incentive is designed to promote positive network effect, i.e. in order for a benefit to be obtained it must be implemented such that the larger the number of firms participating the more benefit for

all the firms and thus the larger the benefit for the public good (Tassey, 1996; Kraut, et. al 1998). It is important to examine incentives, and their correlation to actual usage within the context of network effects.

This paper works to create an understanding of the effect of investments from institutions, generally government or larger monopolistic or duopolistic entities, on actual technology usage. Under the premise that technology usage, rather than just acquisition, is the actual driver for change in an industry. The analysis focuses on the current state of the healthcare system and, more specifically, the impact of incentives from the United States Government, and large hospitals, to physicians for the acquisition and usage of Electronic Health Records (EHR). The paper explores the rationale for the incentives, which is based on driving up the quality of healthcare, ultimately, through the interconnection of these EHR systems. The analysis will attempt to draw a conclusion regarding the effect of these large incentives to determine whether they will be effective in increasing overall actual usage, and ultimately a measure of value-added use of EHR.

Incentive Structures

Incentives are used by large organizations and governments to effect change within an industry (Tassey, 1996). Government incentives can be one of the mechanisms by which particular technologies are created and implemented when the cost or risk is too high for private firms. Government involvement has been instrumental throughout the years, in railroads during the 1800's, the space program in the 1960's, and more recently in the adoption of green energy efficient technologies such as hybrid cars and solar panels.

Government incentives are used when the private market cannot or will not engage in the exchange of goods or services. Normally, this will occur when the cost or risk for production and distribution of these products is abnormally high. However, when this situation arises and the good or service has benefits for others , it is known as a positive externality (Mankiw, 2003). In the presence of positive externalities, where costs are not recovered through the pricing mechanism, government involvement is usually warranted. Market failures are often associated with non-competitive markets, externalities (positive / negative), or public goods (Leyden & Link 1992).

Governments can use a number of different mechanisms to respond to market failures, especially when positive externalities are present, such as in the case of space technology, electric cars, and energy efficient homes. The mechanisms available include tax incentives, subsidies, and price controls (Mankiw, 2003). Responses to market failures depend on the nature of the failure and when positive externalities exists, government should consider subsidies (Tassey, 1996).

While it is evident from the examples listed above that high risk and high cost could be deemed a necessary prerequisite for government involvement, when the cost is not necessarily prohibitive, is there a need for this involvement even in the presence of positive externalities? The private sector will only consider their own particular benefits to the innovation process (Leyden & Link 1992), leading to market failure (Fier, Henerich 2009) regardless of costs. In the case of Information Technology, broadband utilization and penetration has been well documented. Public promotion initiatives was one the key factors in penetration rates of broadband (Miralles, 2006), which has led to

tremendous downstream impact with respect to all of the economic benefits gained from internet usage by organizations and individuals alike. Public policies can also facilitate improvement of products and adaptation of demand (Menanteau & Lefebvre 2000) and thus can change behavior of organizations and individual. The use of incentives, or subsidies, are designed induce behavior and not command it (Fier & Henerich, 2009).

In 2009, as part of the American Recovery and Reinvestment Act (2009), the United States Congress, passed a stimulus bill, which included $19.2 Billion in incentives to physicians and hospitals to adopt Electronic Health Records. More specifically, physicians will be eligible to receive payments of up to $44,000 over a period of 5 years through 2015 for adoption of a certified EHR system. The purpose of this incentive was to increase the adoption rates of EHR systems, which had been less than 20% in 2008. The government estimates that savings of between $81–$162 billion or more annually in healthcare costs would be attained through increased quality of care, better patient outcomes, and a reduction in errors, especially in drug prescriptions and administration (Taylor et. al., 2005).

EHR systems benefit patients, through better care, hospitals, through lower costs and higher efficiency, and the government through lower health care costs and a potentially healthier population (Boyd, et. al., 2010). However, providers don't necessarily see any benefits. Research has found that the low adoption rates are attributable to decreased productivity, prohibitive implementation costs, complexity of systems, and a perception that the system is not necessary (Brailer 2005; Taylor et. al 2005). In 2005, a review of existing literature found major differences between the impact of implemented EHR systems for physicians and nurses. In general while there was some time efficiency gains in documentation for nurses, the same was not true for physicians (Poissant, et. al., 2005). In fact, for physicians, documentation time increased by 17.1%, and bedside Computerized Patient Order Entry (CPOE) work increased the work time over 98.1% per working shift (Poissant, et. al. 2005). The study was performed as part of a review of 23 papers and analyzing the data across each of them.

It is possible that a market failure could be occurring the EHR space, since we have the presence of positive externalities, i.e. benefits to patients, government and even hospitals, but the burden of the costs are imposed on the physicians adopting EHR. Therefore, incentives and subsidies should be considered as a response to the market failure (Tassey, 1996). However, will these incentives actually lead to usage, which in turn lead to the benefits postulated?

The incentive structure of the ARRA (2009) is based on the physician's adoption of the technology and compliance with criteria called 'meaningful use'. Meaningful use is the base criteria established by the Federal Government to receive the incentive payments for EHR adoption. However, long-term benefits of EMR require health care to adopt IT in a purposeful manner (Brailer, 2005). Meaningful use may not be adequate enough to meet the criteria of 'purposeful manner'. The goal is the greatest number of physicians to use the EMR, and not paper, for as many of their daily tasks as possible (Miller & Sim, 2004). Furthermore, effective connectivity, to other providers, is essential to achieving major Health Information Technology benefits, but the barrier to achieving this goal is basic adoption (Taylor, et. al. 2005).

Connectivity to other providers being essential implies a benefit due to network externalities. Previous literature has demonstrated the benefits of breadth of adoption of

interconnected systems. Since the early 1980's Metcalfe's Law has been used as a measure of the value of a network. Originally used to describe telecommunications network, it describes the incremental value of adding a node to a network. Metcalfe's Law has described the value of adding partnership in business (Odlyzko &Tilly, 2005) and more recently in the understanding of social media (Mayfield, 2005). Over the past twenty years, research has both proven and disputed the accuracy and usability of Metcalfe's Law; however, in conjunction with other research such as Reed's Law and Beckstrom's Law, the exploration of the value of networks and its relation to usage is important. As an example, for systems integration, the most important determinant of using a communication system is the total number of people whom potential adopters can reach through it (Kraut, et. al 1998). Thus, people gain more benefit from a communication system if others who are *important to them also use the system* (Kraut, et. al 1998).

Physicians and hospitals are highly dependent on information and require information from various interconnected resources in order to provide quality of care to patients. It has been shown that group technology adoption effected individual technology adoption due to compliance effects (Falk, 1993). Use of communication technology by members of a highly interdependent group increases the value of the technology for all group members and thereby increased their use of the technology (Falk, 1993; Kraut et. al 1998). Therefore, normative and utility characteristics of network externalities will increase adoption of technology. However, there is still a gap in the theoretical review to explain why adoption rates are so low.

Agency Theory

The major cause for the lack of adoption may be interpreted most effectively through principal agent theory. Agency theory focuses on the conflict between two parties when one "hires" another, such that the desires or goals of the principal and agent conflict (Eisenhardt, 1989). The conflict occurs when the principal has no incentive to act (Eisenhardt, 1989; Jaffe & Stavens, 1994).

In the EHR specific case, the physicians become the principals, as they have no incentive to act. Physicians incur all the costs including acquisition, maintenance, training, and costs due to lost productivity. Even in the presence of network externalities, the physicians may not see any benefit to implementing the system.

The agent must therefore be the institution, government or other large organization, offering the incentive. Specifically, the government is attempting to lower its costs associated with healthcare services. However, the source of the conflict occurs because of the nature of the payment system.

Physicians receive payment based on a model called Fee-for-Service (FFS). FFS provides payments to physicians per service rendered to a patient. Therefore, each visit, lab test, or procedure is paid based on rates negotiated by third party payers, or set by government bodies, for Medicare that body is the Center for Medicare and Medicaid Services (CMS).

The conflict therefore becomes very obvious. Physicians incur significant costs in implementing an EHR system, benefits to physicians are debatable, and the overwhelming benefit of EHR systems is the increased quality of care to the patients and

lower payments by the government in the form of elimination of redundant services and errors.

What is the impact of incentives on Information Technology usage in the presence of positive externalities, network externalities and a conflict between principal and agent? The missing piece to explore is to clearly define usage and provide an operational definition for an observational study.

IT Usage

In order to determine the effect of technology on an industry, we must examine its effect at a lower level, either at an organizational level or user level, and then demonstrate that these results, aggregated, significantly impact the industry. Therefore, the first step is to make a determination on how technology impacts firms, which can be effectively done by examining actual usage. However, it is very important to create a more acute and meaningful construct of actual usage.

Davis' seminal paper exploring adoption represents a very critical point for Information Systems research. Much of the academic research leverages the ubiquitous Technology Acceptance Model presented in the paper (Taylor & Todd, 1995). Davis' work is rooted in the Theory of Planned Behavior (TPB) linking attitudes with behavior. TAM provides a parsimonious view of intention to use and has received consistent empirical support for the antecedent attitude attribute of perceived usefulness, while perceived ease of use is mediated by perceived usefulness (Taylor & Todd, 1995).

TAM, however, may not represent a complete model (Agarwahl & Prasad, 1997; Misiolek, et. al. 2002). Additional research comparing different models of usage has consistently validated TAM. Taylor and Todd analyzed competing models of usage by reviewing TAM, Theory of Planned Behavior, and Decomposed Theory of Planned Behavior. Theory of Planned Behavior extends the Theory of Reasoned Action, which incorporates factors of influence external to a user (Taylor & Todd, 1995). The Decomposed Theory of Planned Behavior (DTPB) decomposes beliefs into a multi dimensional construct. Taylor and Todd's findings support the use of the parsimonious TAM model if the intent is to predict usage as a simple construct; however the DTPB may provide more robust understanding of usage behavior.

While TAM continues to be a good measure of system usage, Burton-Jones and Straub (2006) argue for a need to redefine usage and create a better approach to the Usage construct. The authors challenge the usage measures in prior literature on a number of grounds. First, in their review of 48 empirical studies, they found no sound theoretical basis for the usage construct. Furthermore, with the variety of different ways usage is utilized in the research, i.e. dependent variable for some measures and independent variable in other measures, they found the diversity of the usage construct could be problematic, especially in light of the lack theory associated with the usage measures. (Burton-Jones & Straub , 2006).

The review of the literature leads this research to focus on a more quantitative view of usage, as opposed to the qualitative measures in TAM. While it is argued that the non-positivist methods for studying usage are richer than the positivist ones (Orlikowski, 2000), a more quantitative approach is desirable to understand the effects of antecedents on the usage construct. Furthermore, while using the "intention to use"

variable is easily measured in a self-reporting survey, a quantitative approach using data on actual usage can offer a clearer picture of the effects on usage (Devaraj & Kholi 2003) as opposed to the weaker links of self-reported usage (Straub et al. 1995).

Thus usage variables created for this research will be constructed in a manner prescribed by Burton-Jones and Straub (2006). While this model allows for a varying design of the usage construct based on the context, it could be argued that this limits the generalizability of any usage construct. Burton-Jones and Straub(2006) argue that there cannot be a "generally accepted conceptualization of usage" and thus construction of the usage measure(s) are important and should be consistent. This process requires defining the characteristics of the system and assumptions, then defining the structure and function of the usage variables. (Burton-Jones & Straub 2006)

Defining the characteristics includes the definition of a user, subject of the IS artifact, a system, object being used, and a task, function being performed. Specifically, this research focuses on Electronic Health Records, and therefore, the characteristics of the definition usage must be specific to the domain being researched. The user will be defined as the physician, physician's assistant, or nurse as the user. The inclusion of office managers or practice managers should not be included in this construct because the intent of using an EMR are designed around quality of medical care, which means that the provider of the medical care should be the subject of the IS artifact. Secondly, it is clear that the EMR system, the object of the institutional incentives classifies as the system. Finally, the task, the final element of the characteristics of usage, necessary for the Bruton-Jones and Straub contextualization of usage, can be defined as a set of tasks in the EHR.

Tasks used in the construct should be aligned to the "meaningful use" measures necessary for physician's offices to take advantage of the incentive. Meaningful use activities include tasks such as e-prescribing, ability to enter immunizations, and the ability to send information electronically to another system beyond the localized EMR. In order to test a more robust usage measure specifically in this context, allowed by the Burton-Jones and Straub classification, it is suggested that any test of the meaningful use constructs be examined at levels that "meet meaningful use" and exceed "meaningful use". Physicians only interested in meeting the necessary requirements for receiving the incentive may limit the overall benefit, thus usage should be compared at the two levels defined above.

With the usage construct defined, it is necessary to explore performance measures, consistent with Devaraj and Kholi's (2003) research on usage and performance in a hospital setting. Performance should be consistent with the intent of any institutional incentives and thus the paper must examine quality measures. The Federal Government's incentives are designed to improve quality and ultimately reduce costs. Therefore, measures must be found to create this alignment. It is proposed that measures such as errors in e-prescribing, and increases in immunization records should provide solid measures of performances. This research will shy away from financial measures of performance, because the intent of the federal incentives is not centered on the financial performance of physician's offices, but rather the improved quality of care which in turn will reduce the cost of care to the government.

The robustness of the research on the usage construct, enables the creation of a rich measure of usage (Burton Jones & Straub, 2006), which for the purposes of this

research will be split into two separate constructs, usage meeting "meaningful use" and usage that exceeds, or value-added, "meaningful use". This construct allows for performance measures to be compared quantitatively.

Physician's practices do not share the same characteristics. Each practice is different in terms of the location, specialty, pool of patients, etc., and therefore may have different processes in handling patients, which could affect patient outcomes. The differences in organizational characteristics may play a role in adoption of HIT and may impact usage patterns (Hikmet, et. al., 2008). Factors that have been found to impact adoption of HIT in hospitals include size, location (i.e. urban or rural settings, and affiliations (Hikmet, et. al., 2008). Therefore, when studying these physician's practices, it is important to examine the same set of characteristics as they may affect not only the meaningful use, but the value added use construct as well.

Based on the literature review, this research proposal seeks to understand the relationship between institutional incentives and increases in adoption of IT. It has been shown that there is a potential market failure for Electronic Health Record systems, and that based on positive externalities, it is imperative for the government to engage in fixing the market failure.

H1: Institutional Incentives will have a positive effect on basic use

Furthermore, since it has been established that the goal is to achieve the highest level of adoption among physicians, and effective connectivity to other providers is essential to achieving benefits in the system, network externalities are critical to achieving the goals of the incentives.

H2: Increased basic use will have a positive effect on breadth of adoption

The long-term benefits of EHR require health care to adopt IT in a purposeful manner (Brailer, 2005) and furthermore it must be used for as many of the physician's daily tasks as possible. In addition, as more physicians use EHR and are connected to each the more value they will derive from the system.

H3: Increased basic use will have a positive effect on value added use

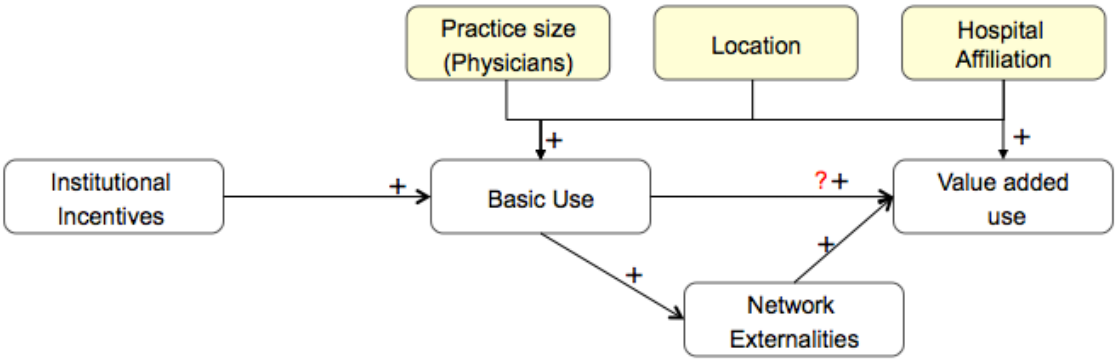
H4: Increase breadth of adoption will positively moderate value added use

Finally organizational factors will play a major role in the adoption of technology, at a basic use level and at a value added use level. Larger physicians practices will have more resources available to them and thus we will be able to leverage the benefits of EHR more readily. In addition, urban-based physician practices will have more connectivity to hospitals and other doctors and thus should benefit from connected EHR systems. Finally, it is posited that physicians who are affiliated with hospitals will also experience higher levels of EHR usage at both a basic and value added use level, since people gain more benefit from a communication system if others who are *important to them also use the system* (Kraut, et. al 1998).

H5a: Larger practices of physicians will see greater basic use
H5b: Larger practices of physicians will see greater value added use

H6a: Urban practices will see greater basic use
H6b: Urban practices will see greater value added use

H7a: Practices with hospital affiliations will see greater basic use
H7b: Practices with hospital affiliations will see greater value added use



Methodology

Field Study & Operational Characteristics

A field study will be conducted to assess the basic and value added usage rates of providers adopting EHR. A survey will be sent to physicians requesting information on their implementation or proposed implementation of an EHR system. Basic demographic and time based information will be collected including size of practice, hospital affiliations, location, and length of implementation or time to implementation. Furthermore, it is hoped that data can be collected directly from the systems to examine the actual usage of the system. If this approach is not feasible then self-reported measures will have to be used.

Operational Constructs

Institutional Incentive: Payments made, or anticipated, by the Department of Health and Human Services (HHS) to providers for meeting meaningful use criteria.
Basic Usage: A subset of key measures defined by Department of Health and Human Services (HHS) for meaningful use needed for incentive payments. Calculated as an aggregated percentage of criteria met (0-100%).
Value Added Usage: Using the same key measures, utilization above meaningful use measures will be assessed Calculated as a percentage above meaningful use (0-100%).

Network Externalities: Number of providers, Health Information Exchanges (HIE), or hospitals, connected, as well as number of records transmitted and received from these entities.

Sample Meaningful Use Criteria

Below is a sample set of criteria of meaningful use measures of EMR that has been defined by HHS in order to be eligible for payment.

1) At least 50% of patients demographic data is stored as structured data
2) At least 50% of patients older than 2 have height, weight and blood pressure recorded.
3) At least 30% of patients have CPOE generated prescriptions
4) At least 40% of patients have CPOE prescriptions electronically transmitted
5) One clinical decision support rule is implemented* (attestation)
6) At least 40% of lab tests are entered into the EMR system

Data Analysis

Once the data is collected, the information will be assessed as part of a series of regression equation to determine the effect on basic usage and value added usage.

Models	
1) Basic Use Model	$Y_1 = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4$
2) Value Added Use Model	$Y_2 = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5$
3) Value Added Use Model with Network Externalities	$Y_2 = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6$

Y_1 – Index of Basic Usage

Y_2 – Index of Value Added Usage

X_1 – Amount of Incentive Received (or eligible)

X_2 – Location (urban/rural) – Dichotomous

X_3 – Size (Physician count) – Continuous

X_4 – Hospital Affiliation (yes/no) – Dichotomous

X_5 – Index of Basic Usage (Meaningful Use)

X_6 – Breadth of network –Continuous variable as a value from 0 – 1. (indexed aggregation)

The first model explores basic usage, as defined by meaningful use criteria. This model takes as its inputs, amount of incentive received (or eligible), location, size, and hospital affiliation. The second model changes the dependent variable, to a value added use construct, and incorporates the meaningful use, or basic usage construct as an independent variable. Finally, the third model, explores the effects of network externalities by including a continuous variable representing an index of the breadth of

the network established by the number of connected physicians and amount of data transmitted.

Additionally, correlation analysis will be conducted against key demographic measures to examine the effect of practice size, location and affiliation. It is expected that affiliation will be significantly high, since many hospital systems are pressuring physicians to adopt EHR and in some cases, such as a large hospital system in the Northeastern U.S., hospital systems are offering their own set of financial incentives for adoption. The R^2 coefficients, and deltas, of each of the models will be analyzed to determine best model. Chronbach's alpha analysis would only be required if the self reported measures were used.

Due to the richness of the data that will be collected, time series analysis could be conducted to examine changes in adoption levels. Further consideration should be given to the possibility of interaction between terms. With additional support alternative techniques could be included such as the consideration of the use of multinomial logistic regression and Structured Equation Modeling.

Expectations

Based on the review of the literature, once the study is completed it is anticipated that institutional incentives will have a positive impact on basic usage, however; it is unclear whether the positive impact of basic usage will lead to a significant increase in value added usage. Should this hold true, it would represent a major obstacle toward the creation of incentives such as subsidies for technology. It would however, not be terribly surprising, since the low adoption rates symptomatic of a lack of perceived usefulness of EHR by providers, which by TAM, would predict that even with incentives, lack of perceived usefulness will be an obstacle to adoption even in the presence of network externalities.

It is anticipated that network effects will have a positive mediating effect on value added usage, however it is expected that this result will be stronger among affiliated physicians, especially those in urban areas, since they will have more contact with other providers and larger pool of patients requiring more coordinated care. Furthermore, larger practices will have more resources available and generally benefit from economies of scale and therefore should more readily adopt EHR and take full advantage of the features within the software.

Limitations

This study is bounded by a few constraints. First, the study only includes EHR systems, which may limit the generalizability of the study. The results may be applicable to other industries; however due to the complex nature of the health care industry this may be problematic. Second, it may be very difficult to obtain the raw data from EHR systems. Providers may not be willing to share data or it may too time consuming to de-identify patient attributes as required under the privacy regulations in the Health Insurance Portability and Privacy Act (HIPAA). In addition, measures of usage must be clear, concise and unambiguous. While the basic use construct can easily be coupled with the meaningful use criteria established as part of the regulation, there may be some

problems with the concepts of value added usage, as a measure above meaningful use. Network effects as a measure of the number of connections to providers and number of inbound and outbound transactions needs clarified. For example, a doctor may have 100 other doctors in their network, but no data is ever sent or received. This could skew the results, but having the large number of doctors available for connection is a key component. Finally, other factors, which could problematic and may have a confounding effect on the regression model include specialty, non-Medicare participants (i.e. Federal Government incentive structure is directed at practices that take Medicare patients and all measures are based on Medicare patients).

Implications for Future Research

Upon completion of this study, other researchers may find the results valuable in other areas. For example, an exploration of the different incentive models (i.e. tax incentives, or non government subsidies) as a predictor of technology usage could prove useful, especially in light of principal agent problems. Since different institutions exert different types of pressure on smaller firms, these results may shed some light on some of the factors of how incentives work to affect the behavior as it relates to usage of Information Technology.

Another interesting area is a closer examination of basic usage and value added usage. The first question that must be addressed is the importance of separating the two constructs. Based on the review of literature, there is an implied need for a value added use construct, but how important this actually is could be a very interesting topic for future research at both a qualitative and quantitative level.

Finally, the Principal-Agent problem with respect to technology usage should be explored more in depth. It would seem that under TAM, technology would not be adopted unless there is a perceived usefulness to the technology; however, this may not represent a complete picture. Solar panels are useful, but aren't broadly adopted for numerous reasons, including uncertainty of the technology and cost. Solar panels do however perform a qualified use, in the sense that they produce energy. It is possible that due to the common nature of technology and the existence of alternatives within Information Technology that perceived usefulness is no longer enough, or will at some point not be enough, to predict adoption. Although this point has been studied extensively, the existence of a principal agent problem within Information Technology may represent an opportunity to explore usefulness, basic usage and value added usage differently.

Conclusion

This research represents a unique opportunity to study incentive structures and its relation adoption of technology. Through well-established guidelines and payment structures, this research can yield insight into the effectiveness of incentive payments on technology adoption. More importantly, the guidelines provided enable the exploration of usage constructs at two different levels, which could prove valuable to other researchers.

Generally, it is expected that incentives will affect some but not all providers and practices; Furthermore, it is expected that different levels of usage will expose problems

with incentive structures even in the presence of varying benefits. The results should raise more questions than provide answers, but the questions will be invaluable in understanding the role of institutions in driving change through the adoption of technology.

References

Agarwal, R., & Prasad, J. (1997). *The role of innovation characteristics and perceived voluntariness in the acceptance of information technologies* - Blackwell Publishing Ltd. doi:- 10.1111/j.1540-5915.1997.tb01322.x

American Recovery and Reinvestment Act of 2009, (2009).

Boyd, A. D., Funk, E. A., Schwartz, S. M., Kaplan, B., & Keenan, G. M. (2010). Top EHR challenges in light of the stimulus. *Journal of Healthcare Informtion Management, 24*(1), 11.

Brailer, D. (2005). Economic perspectives on health information technology. *Business Economics, 40*(3), 6-14. doi:10.2145/20050301

Burton-Jones, A., & Straub D. (2006). Reconceptualizing system usage: An approach and empirical test. *Information Systems Research, 17*(3), 228-246.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340.

Dedrick, J., Gurbaxani, V., & Kraemer, K. L. (2003). Information technology and economic performance: A critical review of the empirical evidence. *ACM Comput.Surv., 35*(1), 1-28. doi:http://doi.acm.org/10.1145/641865.641866

Devaraj, S., & Kohli, R. (2003). Performance impacts of information technology: Is actual usage the missing link? *Management Science, 49*(3), 273-289.

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of Management Review, 14*(1), 57-74.

Fier, A., & Heneric, O. (2009). Public R&D policy: The right turns of the wrong screw? the case of the german biotechnology industry. In M. Keilbach, J. P. Tamvada & D. B. Audretsch (Eds.), *Sustaining entrepreneurship and economic growth* (pp. 147-168) Springer New York. doi:10.1007/978-0-387-78695-7_10

Fulk, J. (1993). Social construction of communication technology. *Academy of Management Journal, 36*(5), 921-950.

Ghosal, V., & Loungani, P. (2000). The differential impact of uncertainty on investment in small and large businesses. *Review of Economics and Statistics, 82*(2), 338-343. doi:10.1162/003465300558722

Gurbaxani, V., & Whang, S. (1991). The impact of information systems on organizations and markets. *Commun.ACM, 34*(1), 59-73. doi:http://doi.acm.org/10.1145/99977.99990

Jaffe, A. B., & Stavins, R. N. (1994). Energy-efficiency investments and public policy. *Energy Journal, 15*(2), 43.

Kraut, R. E., Rice, R. E., Cool, C., & Fish, R. S. (1998). Varieties of social influence: The role of utility and norms in the success of a new communication medium. *Organization Science, 9*(4), 437-453.

Leyden, D., & Link A. (1992). *Government's role in intervention*. Norwell, Ma: Kluwer Academic.

Mayfield, R. (2005). Social network dynamics and participatory politics. In (pp. 116)
Menanteau, P., & Lefebvre, H. (2000). Competing technologies and the diffusion of innovations: The emergence of energy-efficient lamps in the residential sector. *Research Policy, 29*(3), 375-389. doi:DOI: 10.1016/S0048-7333(99)00038-4
Miller, R. H., & Sim, I. (2004). Physicians' use of electronic medical records: Barriers and solutions. *Health Affairs, 23*(2), 116-126. doi:10.1377/hlthaff.23.2.116
Miralles, F. (2006). Efficiency of public promotions policies in the diffusion of broadband networks. an exploratory analysis.
Misiolek, N., Zakaria, N., & Zhang, P. (2002). Trust in Organizational Acceptance of Information Technology: A Conceptual Model and Preliminary Evidence *, 33*
N. Gregory Mankiw. (2003). *Principals of economics* (3rd ed.) South-Western College Pub.
Odlyzko, A., & Tilly, B. (2005). *A refutation of Metcalfe's law and a better estimate for the value of networks and network interconnections.* Unpublished manuscript.
Orlikowski, W. J. (2000). Using technology and constituting structures: A practice lens for studying technology in organizations. *Organization Science, 11*(4), 404-428.
Poissant, L., Pereira, J., Tamblyn, R., & Kawasumi, Y. (2005). The impact of electronic health records on time efficiency of physicians and nurses: A systematic review. *Journal of the American Medical Informatics Association, 12*(5), 505-516. doi:DOI: 10.1197/jamia.M1700
Straub, D., & Limayen, M. (1995). Measuring system usage: Implications for IS theory testing. *Management Science, 41*(8), 1328-1343.
Tassey, G. (1996). Choosing government R&D policies: Tax incentives vs. direct funding. *Review of Industrial Organization, 11*(5), 579-600.
Taylor, R., Bower, A., Girosi, F., Bigelow, J., Fonkych, K., & Hillestad, R. (2005). Promoting health information technology: Is there A case for more-aggressive government action? *Health Affairs, 24*(5), 1234-1245. doi:10.1377/hlthaff.24.5.1234
Taylor, S., & Todd, P. (1995). Assessing IT usage: The role of prior experience. *MIS Quarterly, 19*(4), 561-570.

CAN A ROBOT EVER HAVE SAME CONSCIOUS EXPERIENCE AS A HUMAN?

Jinchang Wang
School of Business, Richard Stockton College of New Jersey
101 Vera King Farris Drive, Galloway, NJ 08205, USA
Jinchang.wang@stockton.edu

ABSTRACT

We argue in this article that a digital machine has an inherent blind spot by showing logically that it will not have same mental experience as a human so far as the self-awareness of an existing person cannot be duplicated. This thesis draws a line between biological humans and digital machines. It makes us rethink the issues such as how far machine intelligence can go, whether robots will eventually dominate humans intellectually, what machines are, and who we are.

Key Words:
Computer Consciousness; Machine Intelligence; Robot; Artificial Intelligence.

Before the first industrial revolution that was driven by mechanical machines and electricity, it was difficult for the people to imagine what a machine *could* do. Nowadays, amid the second industrial revolution driven by computer technology, it is difficult for the people to imagine what a machine *cannot* do. It has been a long-standing contention among scientists on how far intelligent machines can go. Neither side of the contention has so far provided solid arguments for proving this way or the other. We show in this paper that, among the unknowns and uncertainties in the future, there is one thing with certitude: - electronic computers will not have the full range human consciousnesses and mental experience. That is not a hypothesis without solid support. It is a reasoned thesis. We prove it by showing that some human consciousnesses cannot be possessed by a digital robot so far as we do not know how to duplicate the self-consciousness of an existing person.

1. AN OPEN QUESTION

It has been a widely recognized fact that computers are able to do many intelligent things that used to be done only by humans, and that capability is enhancing rapidly. How intelligent a computer can eventually be has been debated for almost sixty years among computer scientists, philosophers, physicists, mathematicians, psychologists, and biologists. Up to now, no one has provided solid arguments to prove this way or the other. How smart a computer will ever be remains an open question.

One side of the contention insists that an electronic machine would, sooner or later, be able to do whatever a human, or a "*meat machine*" by Minsky, is capable of. Here are some assertions by scientists:

"If we're a carbon-based complex, computational, collocation of atoms, and we're conscious, then why wouldn't the same be true for a sufficiently complex silicon-based computer?" – by Gilder & Richards [2]

"In principle, and perhaps in actuality three or four decades from now, it should be possible to transfer a human personality into a robot, thereby extending the person's lifetime by the durability of the machine." – by Bainbridge [1]

"Will these future machines be capable of having spiritual experience? They certainly will claim to. They will claim to be people, and to have the full range of emotional and spiritual experiences that people claim to have." – by Kurzweil [6]

"AI is coming. It is clear we should give conscience to our machines when we can. It also seems quite clear that we will be able to create machines that exceed us in moral as well as intellectual dimensions." – by Hall [3]

On the other hand, many people do not believe that computers can ever be like humans. The strongest among them are dualists who take it for grant that the mind is something separate, and fundamentally different, from the physical things. Many scientists and scholars viewed it unimaginable to have machines sharing all of our intelligence and consciousnesses. Here are some assertions by scientists and philosophers:

"They (computers) are immensely useful devices for simulating brain process. But the simulation of mental states is no more a mental state than the simulation of an explosion is itself an explosion." – by Searle [11]

"I do my best to express, in a dispassionate way, my scientific reasons for disbelieving this perception, and arguing that the conscious minds (of machines) can find no home within our present-day scientific world-view." – by Penrose [9]

"Will emotions be explicitly programmed into a machine? No. That is ridiculous." – by Hofstadter [5]

"Can computers be intelligent? For decades, scientists in the field of artificial intelligence have claimed that computers will be intelligent when they are powerful enough. I don't think so. … Brains and computers do fundamentally different things." – by Hawkins [4]

No one on either side of the contention has so far provided solid reasons to support his/her assertion. They just "feel" that it "should" or "should not". Furthermore, this issue was thought to be neither provable nor un-disprovable. "Someone is bound to ask, can you prove that the computer is not conscious? The answer to this question is: Of course not. I cannot prove that the computer is not conscious, any more than I can prove that the chair I am sitting on is not conscious." [12]

2. CAN A MACHINE EVENTUALLY HAVE ALL HUMAN CONSCIOUSNESSES?

We show that an electronic machine cannot have same consciousnesses as a human by using a counterexample, under the condition that we do not know how to duplicate an existing person's self-consciousness.

Consciousness in this article refers to all mental experience of a person such as thinking, calculation, reasoning, feelings, emotions, intuitions, and faith. A *computer program*, or simply a *program*, refers to instructions to computers in any computer language. *A program is copiable* or *duplicatable* if the instructional statements in the program can be duplicated so that the original and the copy are literally identical and the result of running the copy is indistinguishable from the result of running the original.

Any program for a digital computer is copiable, because, according to the Church-Turing Thesis, it can be converted to a set of equivalent 0-1 codes for a universal Turing Machine [10], and the 0-1 codes on the tape of a Turing Machine are obviously copiable.

A computer is copiable if all the programs in the computer are copiable. All computers as we have at present are copiable since all programs in current computers are copiable.

Self-awareness is a conscious trait "associated with the tendency to reflect on or think about one-self." [8] *Anxiety of death* is comprehension and dread of the mystery / obscurity of death. Self-awareness is necessary to have anxiety of death.

We may have disagreements on exact definitions of self-awareness and anxiety of death. But it would not cause problems in addressing our thesis below. What we need in our argument is just the consensus that both self-awareness and anxiety of death are examples of human consciousnesses.

Anxiety of death defies copying. Suppose robot R is programmed to have self-awareness and anxiety of death. If all the programmed consciousnesses in a robot R are copied to another robot R', then the self-identities of R and R' are same. R and R' share a same "self". Realizing that it has an identical "self" in R, R' would not fear to die since "death" of R' would not result in the ordinary sense of death: - disappearance of the world around itself forever. With the same analysis, robot R would not have anxiety of death either. Therefore, after copying anxiety of death together with self-consciousness from R to R', anxiety of death would disappear on either R or R'. So, we cannot pragmatically obtain a sustained duplication of the consciousness anxiety of death.

An intelligent machine controlled by electronic computers, such as a robot, cannot have sustained consciousness of anxiety of death, as a human, since anxiety of death is copy defiant, under the assumption that no one knows how to duplicate an existing person's 'self'. (Details of the proof will be presented on the conference if this position paper is accepted.) Since anxiety of death is a component of human consciousness, electronic machines cannot have full range human intelligence and consciousness, as stated in the thesis below:

An electronic robot or machine cannot have all human consciousnesses, and so cannot have same mental experience as a human, as far as we do not know how to duplicate an existing person's self-consciousness.

3. DISCUSSIONS

We have reasoned that it is not possible to equip an electronic machine with all human mental experience before we are able to duplicate an existing person's self.

It is "duplicatability" of digital programs that makes a robot incapable to have some consciousness a human has. All digital computer programs are copiable. No man-made machine so far is un-copiable. By our thesis, the machines as we have now will not 'become' humans no matter how complex they will be, as far as we have no idea on how to program or clone a particular person's self-consciousness. In fact, what an un-copiable man-made machine is like, how it works, and how to manufacture such a machine, are still utterly beyond our knowledge.

"Will a robot be one of us?" We now can answer this question with "no, at least before we know how to duplicate an existing person's self-consciousness." We are in the era in which the difference between humans and machines is blurring. Electronic computers have been thought by many to be omnipotent and to surpass humans on every aspect of intelligence and consciousness in a few decades. Our thesis indicates a limit of electronic computers and robots. Even though one could program robots sometime in the future to such a level that they were pretty much human-like, those robots would not be humans because they would miss something we humans have, - our sense on life and death. The electronic robot would not be one of us.

When talking about intelligent machines as a new species created by humans, Kurzweil said, "Evolution has been seen as a billion-year drama that led inexorably to its grandest creation: human intelligence. The emergence in the early twenty-first century of a new form of intelligence (machine intelligence) on Earth that can compete with, and ultimately significantly exceed, human intelligence will be a development of greater import than any of the events that have shaped human history." [7] We can see now that Kurzweil's perspective of machine intelligence is not realistic. The electronic computers as we currently have are not as omnipotent as Kurzweil thought. Even though we do not know how smart electronic computers will be in hundreds and thousands years, they will be incapable of something that humans are capable of, due to the computers' inherent idiosyncrasy, - copiability or duplicatability.

We humbly admit our ignorance about our own consciousness, mentality, and spirit. Is there any piece of our consciousnesses, other than anxiety of death, which may not be realized on electronic machines? Can an existing person's mind be programmed? Can a human's "self" be copied? What is an un-copiable machine like, and how does it work? Is consciousness a "by-product" emerging from sufficiently sophisticate program, as proposed by Hofstadter [5] and some other scientists? If so, how does such "emerged" process occur? Is the emerged process copiable? These are examples of the issues for us to keep reflecting hereafter.

REFERENCES

[1] Bainbridge, W. S. (2004). Progress toward Cyberimmortality. In Ivana Basset ed. Book "The Scientific Conquest of Death – Essays on Infinite Lifespans", Immortality Institute 2004, LibrosEnRed

[2] Gilder, George and Jay Richards (2002), Are we spiritual machines? The beginning of debate. In J. Richards (ed.) Are We Spiritual Machine? – Ray Kurzweil vs. the critics of strong AI. Discovery Institute Press

[3] Hall, J. Storrs (2007). Beyond AI – Creating the conscience of the machine. Prometheus Books, New York

[4] Hawkins, Jeff with Sandra Blakeslee (2004). On Intelligence. Holt Paperback, Times Books / Henry Holt and Company, New York

[5] Hofstadter, Douglas R. (1999). Gödel, Escher, Bach – An eternal golden braid. Basic Books, Inc.

[6] Kurzweil, Ray (2005). The Singularity Is Near – When humans transcend biology. Penguin Books

[7] Kurzweil, Ray (1999). The Age of Spiritual Machines – When computers exceed human intelligence. Penguin Books, Middlesex, England

[8] Encyclopedia of Psychology, Oxford University Press, 2000, Volume 7, p.209

[9] Penrose, Roger (1999). The Emperor's New Mind. 1999, Oxford University Press

[10] Russell, Stuart & Norvig, Peter (2010). Artificial Intelligence – A modern approach. 3rd edition. Prentice Hall, New Jersey.

[11] Searle, John R. (1997). The Mystery of Consciousness. The New York Review of Books, New York, NY

[12] Searle, J. R. (2002). I Married a Computer. In Richard J. (Ed.). Are we spiritual machine? Discovery Institute Press

E-FUNDRAISING - PERSPECTIVES FROM NON PROFIT ORGANIZATIONS

Rhoda C. Joseph, ruj1@psu.edu
Roderick L. Lee, rlee@psu.edu
Pennsylvania State University Harrisburg
School of Business Administration
777 West Harrisburg Pike
Middletown PA 17057

ABSTRACT

E-fundraising (electronic fundraising) refers to the process of soliciting and receiving donations and gifts in a web-based environment. This encompasses the transfer of financial assets via means such as electronic cash, online credit card/debit card transactions, automatic funds transfer, and third parties such as PayPal. Non-profit organizations (NPOs) are uniquely positioned for soliciting and receiving donor funds electronically. The acquisition of funds electronically represents one component of a NPO's fundraising strategy. In this paper we present the case that e-fundraising is a critical strategic undertaking applicable to NPOs given current financial austerity and competition for limited resources. We present a theoretical model founded on the resource-based view (RBV) of the firm, illustrating the primary factors that affect a NPO's e-fundraising strategy. We develop a set of testable hypotheses and examine the implications from both practical and theoretical perspectives in this research-in-progress paper.

KEYWORDS

e-fundraising, e-philanthropy, information technology, non-profit organizations, resource-based view

INTRODUCTION

A NPO also explicitly referred to as a not-for-profit organization is a type of entity that is driven by its social mission and the creation of social value [4] [17]. In the United States NPOs are tax-exempt, cannot generate earnings to benefit an individual or private shareholder, and are classified under the 501c designation. More than thirty different classifications exist under the 501c umbrella, of particular interest to this paper is the 501c (3) category which specifically refers to charitable organizations. Organizations under this umbrella are restricted from activities such as political lobbying but permitted to receive tax deductible contributions from donors in the form of money, volunteer time, and contributed materials.

NPOs are a critical component of the economic, social, and political fabric of the American society. Currently, there are over 1.8 million NPOs operating in the United States. In 2008, their total revenue was $1.9 trillion and total assets were $4.3 trillion [31]. NPOs are operating in an increasingly complex and turbulent environment. For example, the recent economic crisis has resulted in decreased federal funding, decreased charitable contributions, and increased demand

for programs and services [9] [24]. Many NPOs have responded by cutting services and decreasing operating hours. Meanwhile, other NPOs are on the brink of shutting their doors.

It is clear that many NPOs are competing for limited funds from the same pool of donors in order to survive the current operating environment. Ultimately an organization must develop creative strategies in order to enhance its revenue generation. Using information technology to gain a competitive advantage has a long history in the field of strategic management [19] [21] [22]. Researchers have also considered the use of the Internet by NPOs in order to gain a competitive advantage [14].

To examine e-fundraising as a competitive strategy for NPOs we outline the remainder of this paper as follows. In the next section called "Online Fundraising" we examine e-fundraising in the context of charitable giving. The section on online fundraising is followed by a discussion of our proposed theoretical model. There we present a set of testable hypotheses. Next we discuss the proposed methodology and examine implications of the research. We then examine limitations and future directions of this study and present our conclusions.

ONLINE FUNDRAISING

Researchers in the field of online giving or online fundraising have used several different terms including e-philanthropy [11] [23]; e-giving [15]; e-fundraising [28] to classify this concept. Online giving can range from volunteering time to the organization via a web based form to sending actual monetary gifts. As a result of the varying options for online giving, the scope in this study is narrowed to the giving of monetary gifts.

NPOs have a variety of tools and strategies that they use to generate revenue. In addition to government programs and grants, and internal assets and activities, NPOs have three primary sources for generating revenue – individuals, corporations, and foundations [16]. In 2010 individuals, corporations and foundations gave approximately $290.89 billion dollars to NPOs, with a little more than 72% of this giving coming from individuals [16]. As a result of the high percentage of individual giving, NPOs are well served by focusing on this segment and attempting to attract more donations via a web-based environment. From 2009 to 2010 there was a 40% increase in individual online giving to charitable organizations, with the fastest rate of growth seen in small organizations with a database of less than 10,000 email addresses [3]. The decline in corporate philanthropy in the United States [18] is also contributing to a greater focus on smaller donors.

Either the organizational perspective or the individual perspective can be used as the lens to discuss e-fundraising in NPOs. The first perspective looks at the organization itself, and focuses on what makes the organization attractive to specific individual donors. One of the most successful ways to solicit large gifts to NPOs is by initiating and building relationships with current and future donors [5]. NPOs can capitalize on the accessibility of the web, and build relationships with current and future donors through email, online event registration, and volunteer recruitment.

The second perspective looks at the individual and the reasons that motivate giving. Demographic data such as age can shed light on the giving habits of individuals. Based on a self-reporting study conducted in 2010, when asked if they had donated at least $1 to charity during the previous year, 79% of mature adults (born before 1945); 67% of baby boomers (born between 1946-1964); 58% of generation Xers (born between 1965-1980); and 56% of generation Yers (born between 1981-1991) indicated that they had [2]. This study by Bhagat et al. (2010) also showed that there is a steady increase in online contributions to NPOs. The boom of social networking applications such as Facebook and Twitter are creating more points of contact between NPOs and donors. The increased use of social media also creates unique opportunities to attract an individual that is already online, to donate to an organization that he/she is familiar with or believes in.

Indeed, e-fundraising has the potential to increasingly augment and enhance a NPOs overall fundraising strategy. Several studies have focused on philanthropic behavior of individuals with respect to different NPOs. Recent studies have explored a variety of areas such as giving of African American alumni to HBCUs [5] and contribution patterns of different age groups [2]. As e-fundraising is used to increase revenue to the NPO there are potential limitations and hurdles that exist. Lack of trust, lack of accessibility to the technology, and lack of interest in the NPO's mission may inhibit individual giving. Recently the use of kiosk like ATM machines placed in churches for e-giving instead of using collection plates was negatively view by parishioners as cold and disrespectful [15]. As technology changes a traditional environment there will be varying degrees of user resistance. Similarly, if the organization suffers from internal technical, organization, or human factors constraints these may also restrict the NPO's level of interest in e-fundraising.

THEORY AND RESEARCH MODEL

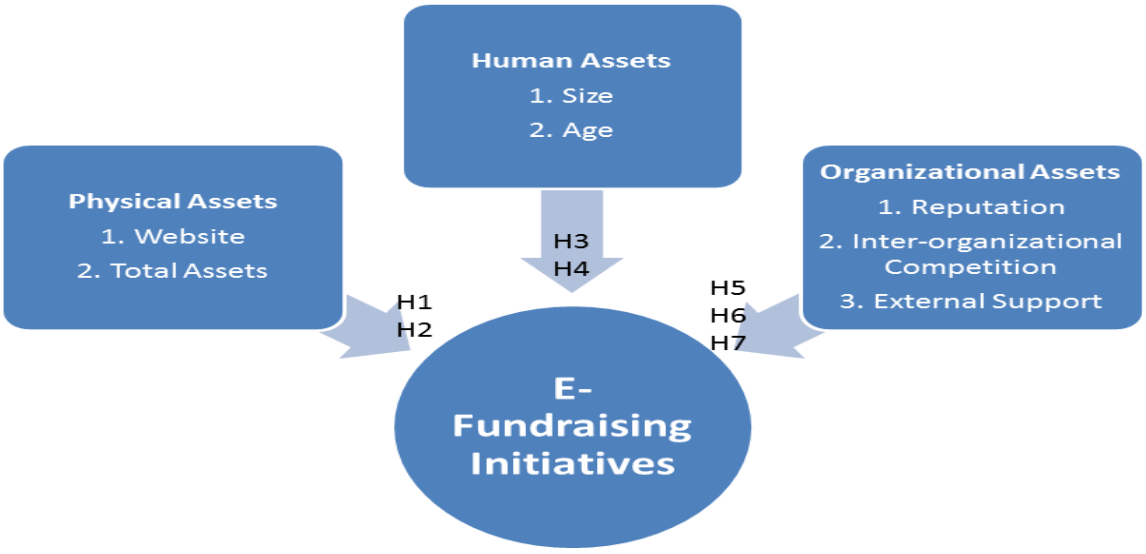
In today's economy many NPOs are competing for limited funds from the same pool of donors. Ultimately an organization must develop creative strategies to differentiate itself from the competition in order to attract and retain donors. For NPOs they must engage in projects that provide an opportunity for the organization to gain a competitive advantage and distinguish itself from the field. Technology can be used by a NPO to gain a competitive advantage through one of four generic strategies (cost leadership, differentiation, cost focus or differentiation focus) and even allow the firm to create new interrelations that did not previously exist [20]. A technology such as the web has spawned the growth of innovative ways to communicate and build relationships. Unfortunately, due to the dependence of volunteer staffing, some NPOs can experience staffing and training challenges that can negatively impact their information technology (IT) portfolio [6]. To gain a competitive advantage the organization must have in-house IT tools and talent or outsource these resources to reap their value.

The resource-based view (RBV) of the firm is one of the most popular and fruitful areas of strategy research [12]. A firm's resources are the assets the firm possesses that enable it to conceive and implement strategic change. The resource-based view (RBV) of the firm states that a firm has three types of resources and capabilities: physical capital resources, human capital resources, and organizational capital resources [1]. First, physical capital resources consist of

financial capital and physical assets such as buildings, and the information technology infrastructure. Next, human capital resources include technical know-how and other knowledge assets such as managerial IT skills. Finally, organizational capital resources include items such as reputation of the firm and the level of external funding. These resources are considered assets that a firm can leverage to exploit strategic opportunities.

We develop a research model based on the RBV. In this model we hypothesize that physical, human and organizational assets directly affect the e-fundraising initiatives undertaken by the NPO. As a result the model consists of seven independent variables and a single dependent variable: e-fundraising initiatives. Below we describe the different variables that make up the model (Figure 1).

Figure 1: E-fundraising Theoretical Model



E-fundraising Initiatives

E-fundraising initiatives are the different online options presented to donors to accommodate electronic payments. The four main categories of e-fundraising initiatives presented in this paper are: 1. printable forms, 2. shopping carts that allow payment by credit and debit cards, 3. third party options such as PayPal and Google check out, and 4. other, which covers internal proprietary systems produced by the NPO. In this study e-fundraising initiatives are collected from the existing web-presence of the organization.

Physical Assets

Physical assets represent the first category of independent variables in the model. Physical assets are divided into two categories – website presence and asset size. A website is a tangible asset that a company possesses. To provide access to e-fundraising options such as shopping carts and third party vendors, it is very useful for a company to have their own website. It is possible that a

NPO can accept gifts indirectly through another organization. However, this scenario is less likely to occur, given the prevalence of web sites today. The evolution of Internet-based technologies such as the Web has created new and exciting opportunities for NPOs to enhance their competitive advantage. Unfortunately, many NPO lack the resources necessary to exploit the strategic potential of the Web [10]. For example, many NPOs lack the financial resources and technical expertise necessary to adopt and implement technologies [6] [25]. While it is clear that technical barriers exist, organizations that have the capacity can harness the value of technology. As a result we expect that NPOs that have a website will be more likely to engage in multiple options for donor giving online. We therefore hypothesize the following.

H1: A website presence is positively associated with e-fundraising initiatives.

An NPO's total assets are reported annually in tax filings (IRS 990). If an organization does not report its assets to the IRS it runs the risk of losing its non-profit status. The asset size of the firm is an indication of the resources that are available to the firm such as physical buildings, land, and equipment. NPOs endowed with a larger number of total assets will have the opportunity to invest in IT applications. Consequently, as the asset value of the firm increases we expect a higher number of e-fundraising initiatives. We therefore hypothesize the following.

H2: The asset size of an organization is positively associated with e-fundraising initiatives.

Human Assets

Human assets represent the human capital of the firm. Human assets contain both tangible components such as the number of employees and intangible components such as managerial skills, training, and in-house technological expertise. Tracking and collecting data on intangible components are not easily quantified. For the human asset, one quantifiable measure is the number of employees or the size of the firm. As firms increase in size the roles of employee and volunteer can be more specialized. Larger firms tend to invest more resources into research and development and innovation [27][j]. A larger organization increases the likelihood of supporting a variety of initiatives, including IT driven innovations such as e-fundraising. We therefore expect that larger NPOs are more likely to engage in e-fundraising initiatives. We therefore hypothesize the following.

H3: The size of an organization is positively associated with e-fundraising initiatives.

The age of an organization is a reflection of its cumulative experience and learning. Therefore organizational age is a proxy for the accumulated competence and increased capabilities that an organization has developed over the years. Older organizations indicate the presence of capabilities to survive under different sets of environmental conditions. As such, older organizations would have a large arsenal of tools to maintain its existence. As part of sustaining future growth organizations with a longer history, are also likely to adopt newer innovative technologies to remain competitive. We expect that organizations with a more persistent presence over time are more likely to have e-fundraising initiates. We therefore hypothesize the following.

H4: The age of an organization is positively associated with e-fundraising initiatives.

Organizational Assets

Organizational assets include the reputation of the organization and external support. Organizational reputation is defined as "a perceptual representation of a company's past actions and future prospects that describe the firm's overall appeal to all its key constituents when compared to other leading rivals" [7].Organizations with esteemed reputations enjoy advantages that cannot be easily imitated by other organizations. The corporate image or brand of a NPO can influence a donor's likelihood of giving to that organization [30]. Organizations with relatively good reputations are more likely to implement e-fundraising initiatives, because of positive perceived reputation. We therefore hypothesize the following.

H5: The reputation of an organization is positively associated with e-fundraising initiatives.

NPOs are competing against other NPOs and in some scenarios are operating in environments where they are also competing against for-profit organizations for resources such as staff, volunteers, and clients. As the competition for limited resources increases, NPOs are more likely to adopt strategies used by for-profit businesses to enhance revenue [29]. Strategies used by for-profits include, mergers, acquisitions, joint ventures and new or spinoff businesses. Given the regulatory constraints of NPOs these options are not always possible but as the number of organizations in a particular sector increases, so too does the level of competition. The use of e-fundraising mechanisms, provide an opportunity for the NPO to differentiate itself or at least remain current with the competition. As the number of firms increase, it typically results in an increase in the amount of competition. We therefore hypothesize the following.

H6: Inter-organizational competition is positively associated with e-fundraising initiatives.

As the number of organizations compete for limited funds, the need for alternative and diverse strategies increases. Survival in competitive environments is based on the establishment of external support networks. Organizations receive external support from foundations, federal programs, charities, and individual donors. In order to receive external support, NPOs have to prove that they are more legitimate than their competitors. One method a firm uses to reduce competition is to develop an innovative product, service or strategy to differentiate itself from the competition [8]. For example, many organizations will include the logo of the United Way on their website to show that they are affiliated with the United Way and receive support from them. We therefore hypothesize the following.

H7: External support is positively associated with e-fundraising initiatives.

PROPOSED RESEARCH METHODOLOGY

This study will use a quantitative analysis of secondary data to test the proposed hypothesis. Even though the majority of prior studies that focused on philanthropy were based on large scale quantitative surveys [5], there are opportunities to use other methods to examine constructs in the e-fundraising domain. To expand the current body of knowledge we employ a method that uses existing data from multiple sources. This study uses empirical data collected from an inventory

of over 1,353 NPOs. We use this data to extract independent data values such as total asset value, size, and age of the organization. We use a second set of data to obtain information regarding the rank or reputation, and external support of the NPOs in its specific service segment. We also use this data set to identify the total number of registered NPOs in that particular category. Lastly, we use the existing website of the NPOs to gather data about all options used to accept financial gifts and donations.

All of the data except for the complete set of values for the dependent variable e-fundraising has already been compiled. Two coders are currently extracting data about e-fundraising from existing NPO websites. We use a tally of the different options to send funds electronically for e-fundraising initiatives. Once all the data on the dependent variable is collected, tests for inter-rater reliability will be conducted. Subsequently, all proposed hypotheses will be tested and analyzed using a statistical software package. This methodology is consistent with the scientific method of an investigative study [26]. All result will then be analyzed and reported.

IMPLICATIONS OF STUDY

This research investigates the strategic use of the web by NPOs to enhance their fundraising efforts. In light of the current economic recession, NPOs are competing for charitable contributions to support their causes. Contemporary research suggests that many technology savvy donors are making donations online. As more individual donors favor the web for donations, it becomes increasingly important for NPOs to adopt a comprehensive e-fundraising strategy. Those that fail to adopt e-fundraising strategies risk becoming obsolete and prone to de-selection from the environment.

This research presents implications for both practice and research. For practice, NPOs will be remiss to not capitalize on the potential of the e-fundraising opportunities available on the web. In this paper we theorize that based on the RBV of the firm, three sets of independent factors affect an NPO's e-fundraising strategy. This can be used as a template for organizations to examine or fine tune current or future e-fundraising initiatives. From a research perspective this study seeks to expand the current literature on technology utilization and strategic e-fundraising by NPOs.

LIMITATIONS AND FUTURE DIRECTIONS

There are two main limitations identified in this study. The first limitation addresses the context of examining e-fundraising. The paper focuses on a single dependent variable of e-fundraising but it does not suggest that the only opportunity for NPOs to generate gifts and donations is through a web-based medium. This study limits its scope and specifically focuses only on the use of web-based methods to solicit and receive funds from donors. E-fundraising is only one part of the overall external funding portfolio implemented by NPOs. E-fundraising is a key component in the organization's arsenal of fund raising tools because it is effective and cost restrictive, but it should be incorporated with other traditional fundraising options such as radio, direct mail, telephone, in-person visits, and other forms of donor communication [11]. This study is not

suggesting that e-fundraising replace other effective strategies, but instead be included in conjunction with others.

The second limitation of this study addresses the components of the theoretical model. It is possible that since our theory is informed by the RBV of the firm, there are other independent variables outside the scope of the RBV that are not captured. One such example could be the amount of litigation that the NPO has been exposed to. NPOs are susceptible to security breaches but unlike for-profit businesses, they may not have the necessary resources to counteract such attacks [13]. The existence of online data breaches may restrict the growth of web-based services that the NPO provides. However, we have refrained from speculating on other factors that may be present, because this may already be captured in the reputation of the organization. In an alternative model, one study showed that accessibility, accountability, education, interaction, and empowerment were all positively correlated to e-philanthropy [23]. That study however focused only on characteristics of the NPO's website. In our current study we present a more comprehensive model, and look beyond only the factors discernible from the website. Exploring a different set of constructs can be the subject of future research projects.

CONCLUSION

This study uses the resource based view of the firm to examine the factors that affect a not-for-profit's e-fundraising strategy. It is important to examine this construct of e-fundraising since NPOs represent the third largest sector of the U.S. economy after for-profit businesses and the government. Understanding the strategic drivers of e-fundraising can be used to identify best practices and opportunities for growth in the future. Currently the U.S. economy is facing financial and economic challenges, and these challenges such as unemployment, underemployment and wage stagnation have a direct impact on the not-for-profit sector. As individuals are more careful and precise with how they allocate their limited financial assets, NPOs need effective tools to attract and maintain future and current donors. This research can provide insights for effective e-fundraising strategies for researchers, practitioners, and any other entities that are committed to and support the missions of NPOs.

REFERENCES

[1] Barney, J.B. *Firm resources and sustained competitive advantage*, Journal of Management 17, 1991, pp. 99 - 120.

[2] Bhagat, V., P. Loeb, and M. Rovner, *The Next Generation of American Giving*, in *A study on the contrasting charitable habits of Generation Y, Generation X, Baby Boomers and Matures*, Convio, 2010, pp. 16.

[3] Bhagat, V., D. McCarthy, and B. Snyder, *The Convio Online Marketing Nonprofit Benchmark Index Study*, Convio ed., 2011, pp. 48.

[4] Bryson, J.M. *Strategic planning for public and nonprofit organizations: A guide to strengthening and sustaining organizational achievement* 3ed, Jossey-Bass, San Francisco, CA, 2004.

[5] Drezner, N.D. *Why give?: Exploring social exchange and organization identification theories in the promotion of philanthropic behaviors of African-American millennials at private-HBCUs*, International Journal of Educational Advancement 9, 2009, pp. 147–165.
[6] Evans, S.H., and P. Clarke, *Training Volunteers to Run Information Technologies*, Nonprofit and Voluntary Sector Quarterly 39, 2010, pp. 524-535.
[7] Fombrun, C. *Reputation: Realizing value from the corporate image*, ed, Harvard Business School Press, Boston, MA, 1996.
[8] Gilbert, R.J. *Competition and Innovation*, in *University of California, Berkeley*, University of California, Berkeley, 2006, pp. 1-31.
[9] GuideStar, *The effect of the economy on the nonprofit sector, March - May 2009*, July 30.
[10] D. Hackler, and G.D. Saxton, *The strategic use of information technology by nonprofit organizations: Increasing capacity and untapped potential*, Public Administration Review 67, 2007, pp. 474 - 487.
[11] Hart, T.R. *ePhilanthropy: Using the Internet to build support*, International Journal of Nonprofit and Voluntary Sector Marketing 7, 2002, pp. 353-360. Available at http://onlinelibrary.wiley.com/doi/10.1002/nvsm.192/pdf.
[12] Hoopes, D.G., T.L. Madsen, and G. Walker, *Guest editors' introduction to the special issue: Why is there a resource-based view? Toward a theory of competitive heterogeneity*, Strategic Management Journal 24, 2003, pp. 889 - 902.
[13] Kolb, N., and F. Abdullah, *Developing an Information Security Awareness Program for a Non-Profit Organization*, International Management Review 5, 2009, pp. 103-107.
[14] Lee, T.E., J.Q. Chen, and R. Zhang, *Utilizing the Internet as a competitive tool for non-profit organizations*, Journal of Computer Information Systems 41, 2001, pp. 26 - 31.
[15] McBride, J. *Locals Weigh in on E-giving, Donation Kiosks*, McClatchy - Tribune Business News 2008, pp. 1. Available at http://ezaccess.libraries.psu.edu/login?url=http://search.proquest.com/docview/465518474?accountid=13158.
[16]NCCS, 2011, "Quick Facts about Nonprofits," *National Center for Charitable Statistics, http://nccs.urban.org/statistics/quickfacts.cfm*
[17] Moore, M.H. *Managing for value: Organizational strategy in for-profit, nonprofit, and governmental organizations*, Nonprofit and Voluntary Sector Quarterly 29, 2000, pp. 183 - 208
[18] Porter,M. E., and M.R. Kramer, *The Competitive Advantage of Corporate Philanthrophy*, Harvard Business Review, 2002, pp. 5-16.
[19] Porter, M.E., *Competitive advantage: Creating and sustaining superior performance* ed, Free Press New York, NY 1985.
[20] ---, *Technology and Competitive Advantage*, Journal of Business Strategy 5, 1985, pp. 60-78.
[21] ---, *Competitive strategy: Techniques for analyzing industries and competitors*, ed, Free Press, New York, NY, 1998.
[22] M.E. Porter, and V.E. Millar, *How information gives you competitive advantage*, Harvard Business Review 63, 1985, pp. 149 - 160.
[23] Sargeant, A., D.C. West, and E. Jay, *The Relational Determinants of Nonprofit Web Site Fundraising Effectiveness An Exploratory Study*, Nonprofit Management & Leadership 18, 2007, pp. 141-156.

[24] Saxton, G.D., and C. Guo, *Accountability online: Understanding the Web-based accountability practices of nonprofit organizations*, Nonprofit and Voluntary Sector Quarterly 40, 2011, pp. 270 - 295.
[25] Schneider, J.A. *Small, minority-based nonprofits in the information age*, Nonprofit Management & Leadership 13, 2003, pp. 383 - 399.
[26] Sekaran, U. *Research Methods for Business: A Skill-Building Approach,* ed, John Wiley & Sons, New York, 2000.
[27] Shefer, D., and A. Frenkel, *R&D, Firm Size and Innovation: An Empirical Analysis*, Technovation 25, 2005, pp. 25-32
[28] Stephens, J.L. and R.W. Clouse, *Impact Of E-Fundraising for Nonprofit Organizations: An Entrepreneurial Approach*, in *Allied Academies International Conference*, Proceedings of the Academy of Entrepreneurship, Nashville, TN, 2002, pp. 49-58.
[29] Tuckman, H.P. *Competition, commercialization, and the evolution of nonprofit organizational structures*, in *To profit or not to profit: the commercial transformation of the nonprofit* B.A. Weisbrod ed., Cambridge University Press, 1998.
[30] Venable, B.T., et al., *The Role of Brand Personality in Charitable Giving: An Assessment and Validation*, Journal of the Academy of Marketing Science 33, 2005.
[31] Wing, K.T., K.L. Roeger, and T.H. Pollak, *The nonprofit sector in brief: Public charities, giving, and volunteering, 2009*, The Urban Institute, Washington, D.C., 2010.

THE QUEST FOR BUSINESS INTELLIGENCE

Santosh S. Venkatraman; Department of Business Information Systems ; College of Business; Tennessee State University; Tel: (615) 963-7315; svenkatraman@tnstate.edu

Alicia M. Brooks, Senior Analyst, Prevention & Shrink Improvement; Dollar General Corporation; 100 Mission Ridge; Goodlettsville, TN 37072; abrooks@dollargeneral.com

ABSTRACT

This paper provides an insight into how modern organizations use business intelligence (BI) systems to make better decisions, and gain a strategic advantage over their competition. We discuss the nature of business intelligence tools along with popular features such as dashboards and scorecards. The hot trends in BI such as data visualization, cloud computing, spatial visualization, and predictive analytics are also identified and described. This research will prove beneficial to practitioners, as well as academicians who want to pursue research in the business intelligence field.

INTRODUCTION

With the proliferation of information technology (IT), businesses are often overwhelmed with the sheer quantity of data, as every business process can potentially accumulate millions of data points on a daily basis. Data and information such as the turnover rate in a particular department, sales figures, shipping quantities, customer loyalty information, product specifications, or even gasoline prices and weather information are all measured and stored by corporations around the clock. Wal-Mart, for example, handles more than 1 million customer transactions every hour and generates more than 2.5 petabytes of data storage, which is equivalent to 167 times the books in America's Library of Congress" [3].

With such vast amounts of data that are amassed and available, it is imperative to get the correct information to the correct person at the right time in order for business decisions to be made successfully. A natural dilemma is how businesses can make sense of all the data without wasting time and resources as the amount of data captured continues to soar? Many corporations are turning to enterprise data warehousing and business intelligence tools in order to assist with the otherwise overwhelming task of organizing the data and making informed decisions. Even though data warehouses have been used effectively by businesses for many years for decision support and planning [6], the additional deployment of newer business intelligence tools makes decision making even more effective and efficient.

The purpose of this paper is to provide an insight into how modern organizations use business intelligence (BI) systems to make better decisions, and gain a strategic advantage over their competition. The next section discusses the nature of business intelligence tools along with their popular features such as dashboards and scorecards. Next, we describe the hot trends in BI such as data visualization, cloud computing, spatial visualization, and predictive analytics. We summarize and conclude the paper in the final section.

NATURE OF BUSINESS INTELLIGENCE

Business intelligence (BI) can be defined as timely, accurate, high-value, and actionable business insights, and the work processes and technologies used to obtain those [16]. Business intelligence environments consist of the hardware, software and organizational support for business intelligence activity that enables knowledge workers to access, analyze and manipulate data [10]. Many BI software systems are installed in conjunction with data warehouses as the combination of the technologies is unparalleled in terms of efficiency, data management, and ease of use. BI platforms allow users to convert raw data into visually appealing and dynamic graphs and reports through the use of a graphical user interface. With business intelligence applications, managers do not have to be advanced statisticians in order to meaningfully understand the data available. In fact, the concept of business intelligence is intended to be so user-friendly that in ideal conditions the end user is not even aware of the fact that they are employing BI tools [9]. Figure 1 [11] illustrates the wide variety of BI tools, such as dashboards and scorecards, that can be used by various end-user types, such as front-line workers and Managers. Some of the major BI vendors include Revolution Analytics, Cognos, SAP, MicroStrategy, Oracle, IBM, Teradata, SAS, SPSS, and Microsoft.

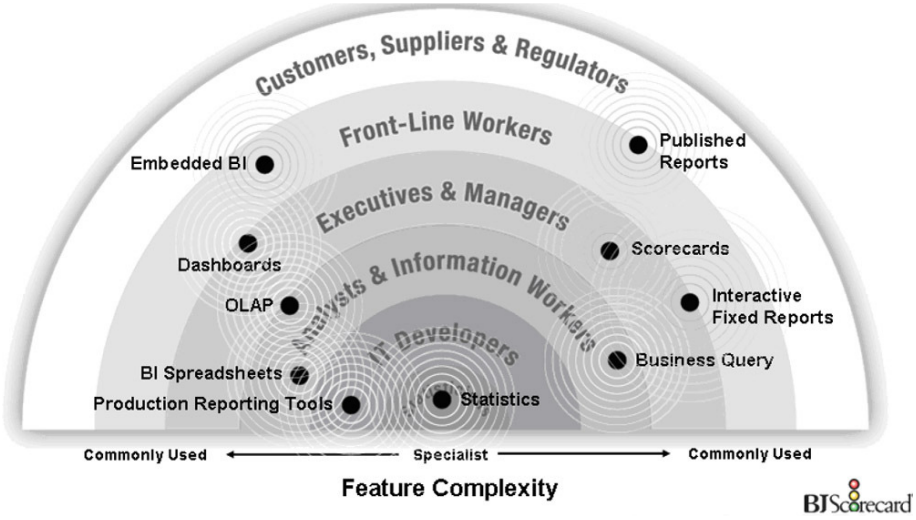


FIGURE 1: BI Tools and User Types

The real attractiveness of incorporating a BI platform is that it can be completely customized and built to meet the needs of the organization. There are a wide variety of tools and features available in a BI system as described below:

Data Visualization is arguably the most widely-used tool within BI platforms today. Data visualization transforms raw data into visually appealing graphical depictions that communicate almost instantaneously with business end users. Organizations today are experiencing an undeniable cultural shift in which business users are wielding greater influence over budgetary and purchasing decisions that were previously dictated by the IT department [2].

It stands to reason that business users will most likely place emphasis on different features or functionalities than IT users, and those features will tend to drive BI purchases. Business users value data visualization because it allows them to promptly identify issues, patterns, outliers, or trends and take necessary action without having to manually search through text or data tables [4]. Users can see the big picture, able to save time and minimize potential human error. Figure 2 below shows a graphical data visualization of retail sales.

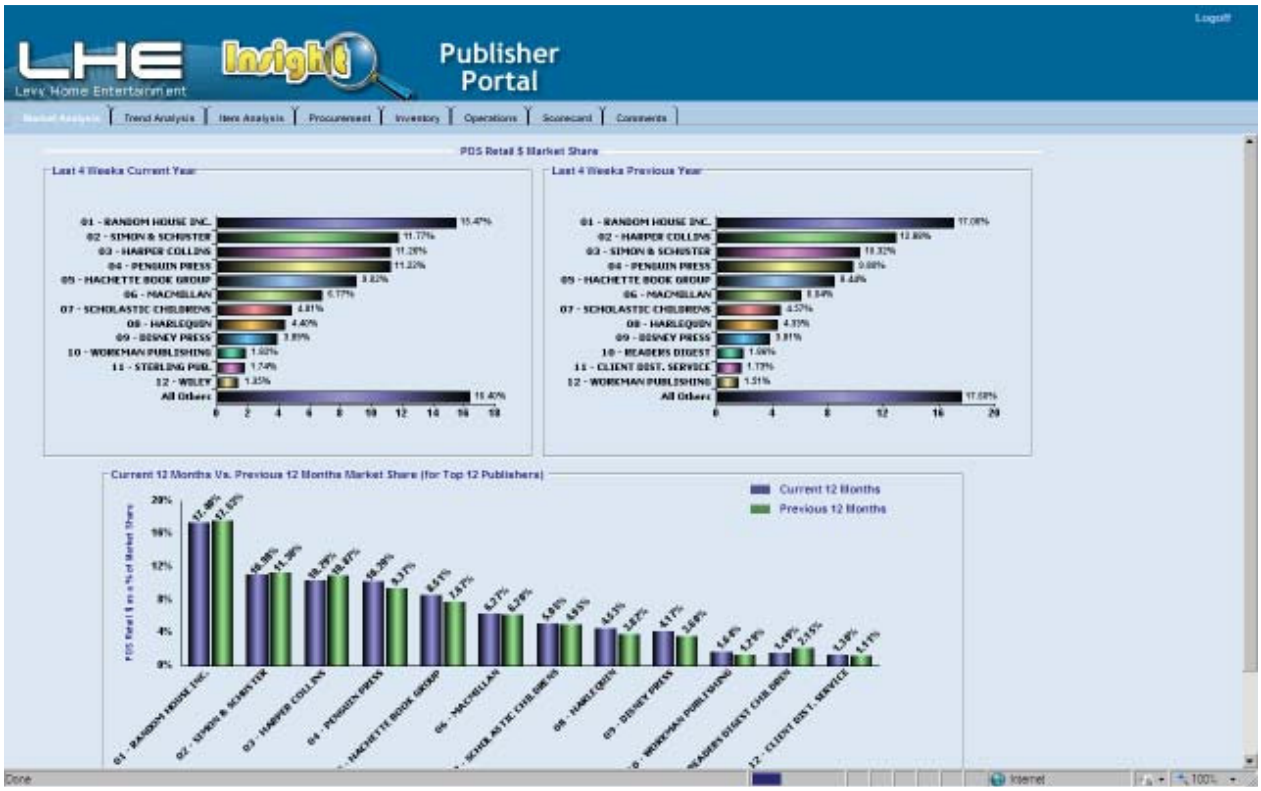


FIGURE 2: Visual Reporting In Retail Sales

Dashboards and Scorecards are also a popular feature of many BI tools. Regardless of industry, organizations are adopting dashboards and scorecards to measure key performance indicators (KPIs) and drive results in focused business areas. Key performance indicators are specified metrics whose measurements provide insight into key organizational operations, initiatives, or processes and help management make better decisions.

The Data Warehouse Institute (TDWI) recently surveyed [4] over 200 professionals regarding the use of visual reporting and analysis for its best practices report, and an overwhelming number of respondents indicated that dashboards are the favored instrument for visual reporting. Many dashboards allow a "one-stop shop" for users by displaying any KPIs or data points that are driving the focus of the organization; the user can quickly access the information that he or she needs to take appropriate action and make the right decisions for his or her business unit. The best dashboards and scorecards enable users to drill down one or more levels to view more detailed information about a metric. In essence, a dashboard is a visual exception report,

highlighting performance anomalies using visualization techniques [4]. One popular option in visual reporting is the use of stoplight formatting. Based on the concept of a traffic light, values are highlighted in red (indicating an unfavorable result) or green (indicating a favorable result). Figure 3 shows a sample dashboard featuring KPIs, charts, filters, and stoplight formatting for a manufacturing application [14].

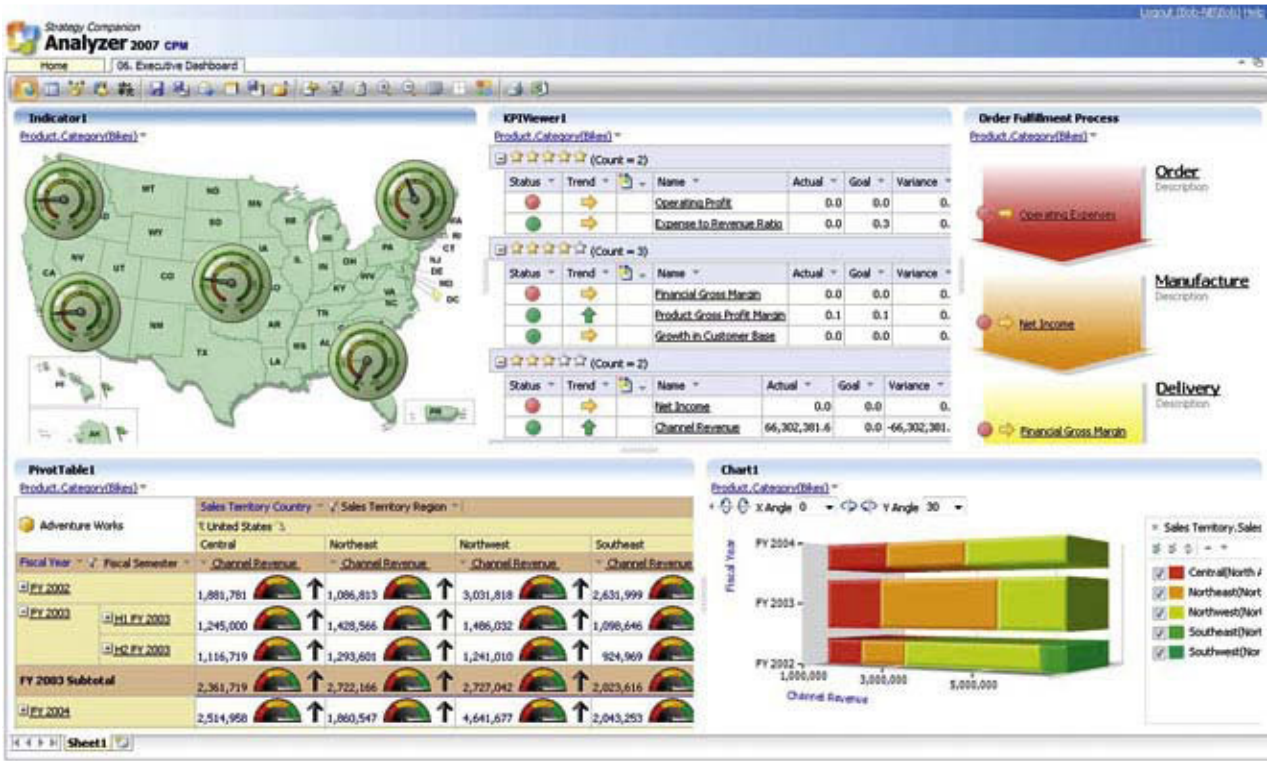


FIGURE 3: Manufacturing Application Dashboard

Dashboard designs are as unique as the organization for which they are employed. While most dashboards tend to share certain design characteristics, there are many variations to the design layouts and possibilities. Most dashboards arrange a series of related charts in a grid template, and use multiple tabs or radio buttons to segment charts by category. They also usually display filters above or beside the charts they apply to, as well as links to related dashboards or reports. The best dashboards display summary data graphically so it can be consumed at a glance and then provide access to any detailed information within a few mouse clicks [4]. The use of features such as drill-down links and filters allow business users to view high level summary data and then dive into details for those metrics that draw attention, rather than having to manually search for the granular data. The high level of interactivity in dashboards is appealing to many business users as they make business decisions. Figure 4 [6] depicts an alternative dashboard layout based on daily and monthly KPIs measuring customer satisfaction and performance within the hospitality industry.

FIGURE 4: Sample Performance Dashboard

Data Discovery is a class of data visualization application that is gaining popularity among business users. Data discovery software offers an alternative to traditional BI platforms and appeal to end users because of the ease of use and implementation. Unlike traditional BI platforms which are bought and supported by IT departments, data discovery tools are designed to support smaller groups of users and focus more on analysis than reporting. They are typically more expensive when considering the cost-per-user, but are popular any way because of the ease-of-use and great benefits. A potential risk to the organization is that business users are willing to use data discovery applications and create information silos within their departments. This software, therefore, tends to work best in smaller organizations, and although it is possible to adopt data discovery as a stand-alone technology, best practices suggest that this technology be used in conjunction with a data warehouse. Specialized vendors like Tableau, QlikTech, or Spotfire are experiencing success with data discovery software options, but larger vendors such as Microsoft, SAP, IBM, and Information Builders are reacting with their own versions to increase competition in this market segment [2].

CURRENT TRENDS IN BUSINESS INTELLIGENCE (BI)

With increasingly sophisticated software, fast computers and communications networks, the area of business intelligence is evolving rapidly to meet user needs. We next describe some of the major trends that are shaping BI.

BI & Real-Time Analytics

In this fast paced world of global business and blazing networks, it is indeed valuable to know the current status of your organization at all times. Why make a decision based off of last week's data when this week's information is available instead? Active data warehousing is based off the

concept of nightly batch updates that insert the activity of the day into the data warehouse appropriately. This daily process results in records that are relevant, timely, and actionable – but still not "real-time" or even "near-real-time." Modern managers, however, are experiencing a cultural shift and require that data and information be available as it is happening. This need for real-time analytics is extending far beyond its traditional users, such as airlines or emergency response units [5].

Even retailers and television networks are now implementing software to deliver information in as near-real-time as possible. One example of a successful near-real-time analytics adoption is Haggen, Inc., a small retail grocery chain based out of Washington State. Traditionally, grocery items tend to have low margins and perish/expire relatively quickly; therefore data mining used to be conducted long after the actual sale had taken place. However, when management at Haggen realized that an empty shelf space was the equivalent of lost sales, they decided to opt toward a more near-real-time approach to inventory management. The grocer began using software that pulls information from a data warehouse using continual online transaction processing rather than daily batch jobs, allowing managers to focus on making timely decisions. Before the incorporation of real-time analytics, managers did not have access to the previous day's activities until 9 a.m., which could impair effective decision making for a grocer that is preparing items in the bakery or deli. According to Haggen's CIO Harrison Lewis, they now get visibility throughout the day; and have information that is just 10 to 15 minutes old. Since Haggen deals mainly with products that are made in the store, having the ability to respond faster gives them the ability to take advantage of a good situation or minimize the impact of a bad one [5].

Cloud-Based Business Intelligence.

Cloud computing is another technology trend that has gained popularity in recent years as it allows users to access IT resources like data storage or application services such as business intelligence or CRM. There is no need for the user to purchase, install or maintain expensive hardware and software. The IT resources are typically accessed from the "cloud" (essentially a metaphor for cyberspace) using a web-browser, or a lightweight "app" installed on a computer or a mobile device such as laptops, tablet computers or smart-phones. Figure 5 illustrates the typical structure of cloud computing [8].

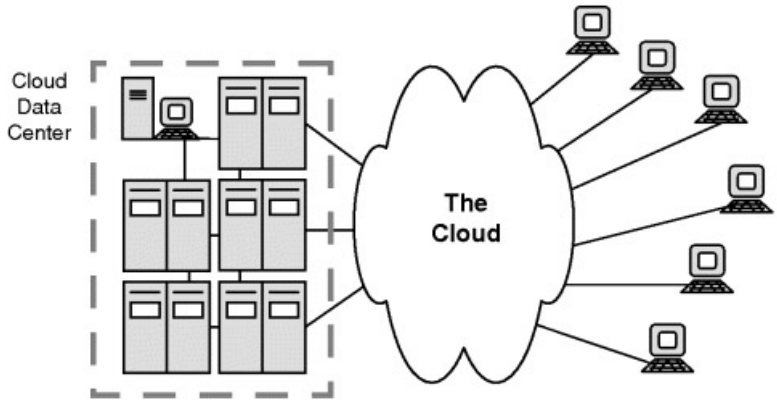


FIGURE 5: Cloud Computing

When it comes to cloud computing in BI environments, most organizations typically contract vendors, such as 1010data, Elastra, Rightscale, Oracle, and IBM who specialize in cloud computing. All maintenance, hardware, and software issues will be the responsibility of these third party vendors, and the organization can instead focus its resources on its core competencies. Of course, the organization has to pay the cloud computing vendor periodically for the "cloud-based services," but they only pay for what they use, much like for a utility such as electricity.

On-demand BI offerings with utility based licensing have also piqued the interest of a growing number of users as a means of cutting costs and boosting profits. Business Intelligence (BI) and Analytics in the cloud is a fast growing market, and IDC expects a compounded annual growth rate of about 22% through 2013 [7]. This growth is primarily driven by two kinds of cloud applications. The first is a purpose-specific analytics-driven application for business processes such as financial planning, cost optimization, and inventory analysis. The other is a self-service horizontal analytics application that allows the customers to analyze data and create, embed, and share analysis and visualizations.

BI & Spatial Visualization

When technology combines the power of BI analytics with global positioning systems (GPS), users are not only able to see what is happening – but also where it is happening. Figure 6 depicts how the number of customers for a particular business stacks up in different cities in Texas, and it also shows the distance and direction of the customers from the factory in Abilene, TX [16].

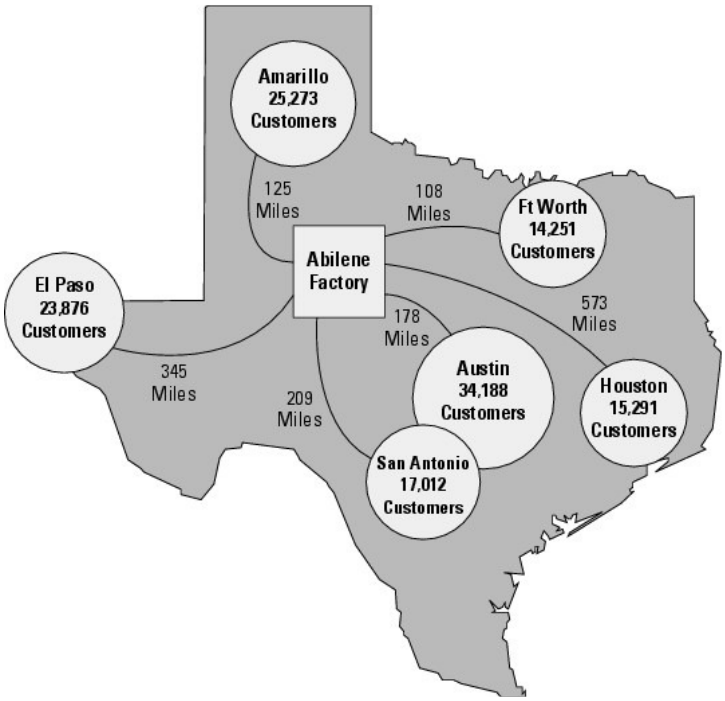


FIGURE 6: Customer Data in Spatial Visualization

As another example, consider a manager viewing a map that shows negative year over year comparison sales for a specific location for a particular fiscal period. The reason for the

decreased sales could probably be very apparent if the map also showed that the region just experienced a major natural disaster. This visual recognition will help the manager make intelligent business decisions quickly, rather than wasting time trying to "guess" the possible reasons for the sales decline. A final example of successful spatial visualization use would be competitor analysis. By mapping new, remodeled, or relocated competitors, an organization could quickly identify locations that will potentially be impacted by the incoming competitor. This knowledge will allow the organization time to proactively implement an effective strategy within those areas and minimize negative impact.

BI giant Oracle has incorporated spatial visualization into its popular Oracle Business Intelligence 11g platform. Figure 7 illustrates a spatial visualization that measures units and variable sales of a particular set of products in a manufacturing organization [12].

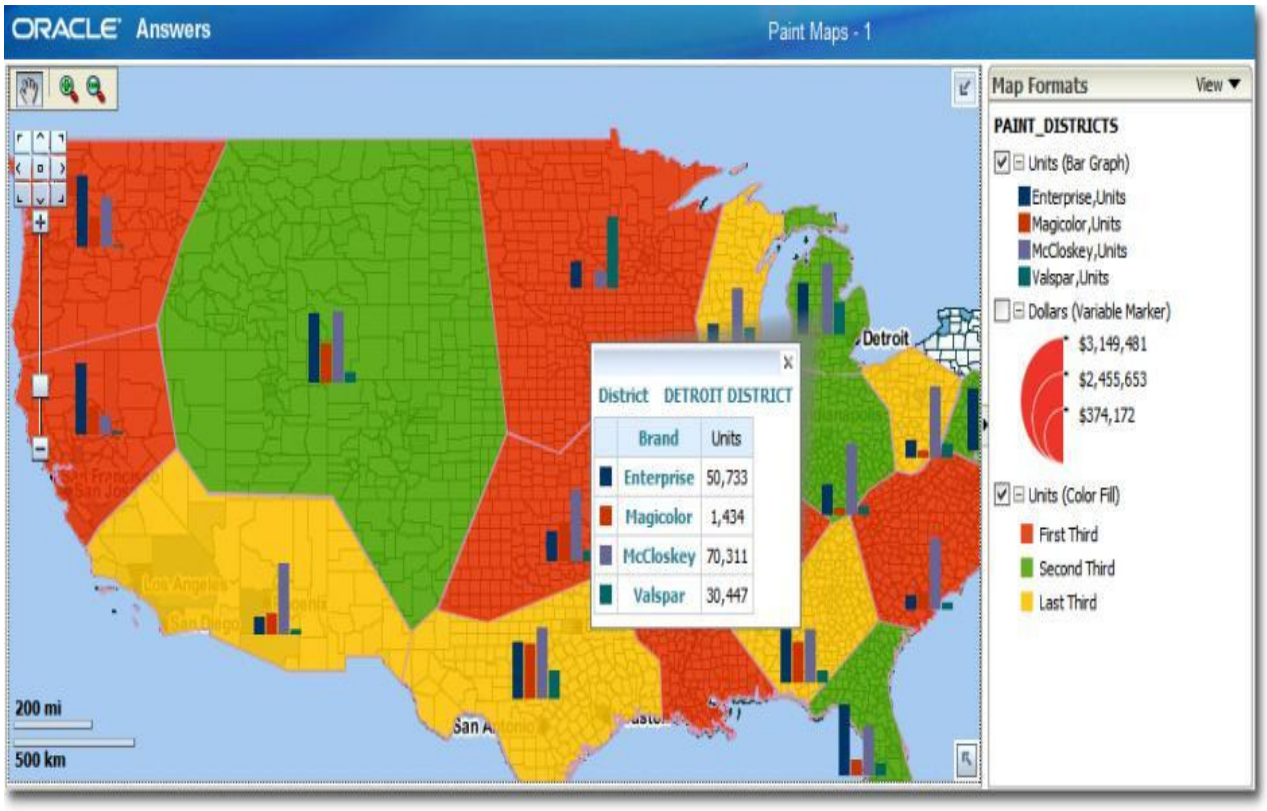


FIGURE 7: Spatial Visualization using Oracle BI 11g

BI & Predictive Analytics

Predictive analytics is another very hot trend in the BI field that attempts to predict the future. At a time when business users can easily capture and analyze previous performance, as well as events as they happen in real-time, predictive analytics has become the new frontier of the analytical realm. Business users not only want to understand what has occurred or is occurring currently, but also desire to know what is in store for the future. The use of predictive analytical tools and systems attempts to make this possible in an effective and efficient manner. These

systems delve into an organization's data to recognize patterns, trends, outliers, commonalities, or any other relationships evident in the dataset. Using sophisticated algorithms and complex modeling methodologies, these systems go a step beyond identifying the patterns and actually predict what behaviors might occur in the future based on historical data.

The implications that this technology will have on organizations are immeasurable. From a marketing perspective, an e-commerce site could literally calculate the next mouse click or key stroke of a customer based on trends, and strategically position advertising according to the expectations. The realm of public services could find value in these applications as well. What if local law enforcement could predict when and where crimes will occur, and accordingly allocate officers in those areas at those times? If business intelligence is largely based on the ability to use what has happened or is currently happening to make the best possible business decisions, then predictive analyses could very well be the frontrunner of the BI user-community.

SUMMARY AND CONCLUSION

It is essential for modern managers to have access to relevant information at the right time and format so that they can make effective decisions efficiently. With the incessant collection, storage and processing of large volumes of data, it is easy for many managers to get caught up in "analysis paralysis," and subsequently make sub-optimal decisions. As seen in this paper, a Business Intelligence (BI) system can be a very effective tool in the managers' arsenal so that the right information is delivered to the right user at the right time in the right format.

In a world where the volume of digital material growing approximately by ten times every five years, businesses which adopt appropriate technologies and intelligence systems will be more successful than their non-technological counterparts. Data are becoming the new raw material of business, and serves as an economic input almost on a par with capital and labor [3]. The future ultimately will likely belong to business managers who can maximize the profitability of their organizations by efficiently and effectively using this new raw material – data. With a plethora of BI tools such as dashboards, scorecards, spatial visualization and predictive analysis, business managers who leverage BI systems to their advantage can expect to be very successful. This paper will prove beneficial to practitioners, as well as academicians who want to pursue research in the business intelligence area. It will prove beneficial if future research in BI identify critical success factors for BI implementations, and disseminate best practices in this fast growing and exciting field.

REFERENCES

[1] Atre, S. (2011, January 27). DW Basics. Retrieved January 29, 2011, from Information Management: http://www.information-management.com/infodirect/2009_189/data_warehouse_BI_ETL_analytics_metadata-10019591-1.html

[2] Brunelli, M. (2011, February 11). *Gartner Magic Quadrant: Data discovery vendors take on BI heavyweights*. Retrieved April 6, 2011, from SearchBusinessAnalytics.com: http://searchbusinessanalytics.techtarget.com/news/2240032062/Gartner-Magic-Quadrant-Data-discovery-vendors-take-on-BI-heavyweights

[3] Cukier, K. (2010, February 25). *Data, Data Everywhere. The Economist*. http://www.economist.com/node/15557443

[4] Eckerson, W., & Hammond, M. (2011, January 12). *Visual Reporting and Analysis: Seeing is Knowing.* Retrieved February 3, 2011, from TDWI: http://tdwi.org/webcasts/2011/01/visual-reporting-and-analysis-seeing-is-knowing.aspx?sc_lang=en

[5] Fogarty, K. (2008, May 2). *How to Get Real-Time Analytics from a Data Warehouse*. Retrieved February 24, 2011, from BaselineMag.com: http://www.baselinemag.com/c/a/Enterprise-Apps/How-to-Get-RealTime-Analytics-from-a-Data-Warehouse/

[6] Getz, A. (2006, December). *DW Basics.* Retrieved January 14, 2011, from Information Management: http://www.information-management.com/white_papers/2292752-1.html?portal=dw_basics

[7] Gopalakrishnan K. (2011, May 27). Business Intelligence and Cloud Computing. Retrieved from http://blogs.hexaware.com/business-intelligence/business-intelligence-and-cloud-computing/

[8] Hugos, M., & Hulitzky, D. (2011). *Business in the Cloud: What Every Business Needs to Know About Cloud Computing.* John Wiley & Sons.

[9] Imhoff, C. (2009, September 23). *Data Warehousing & Analytics Library.* Retrieved March 17, 2011, from IBM: http://www-01.ibm.com/software/data/infosphere/data-warehousing/library.html

[10] Mosley, M. (Ed.). (2008). *The DAMA Dictionary of Data Management* (1st Edition ed.). Technics Publications.

[11] Mosley, M. (Ed.). (2010). *The DAMA Guide to the Data Management Body of Knowledge* (1st ed.). Technics Publications.

[12] Oracle Corporation. (2010). *Introduction to Geospatial Visualization in Oracle Business Intelligence Enterprise Edition, 11g.* Retrieved April 2011, from Oracle.com: http://download.oracle.com/otndocs/products/mapviewer/pdf/mv11g_spatialvis_inobiee.pdf

[13] Quinn, K. (2011). *Improving Operational Efficiency in Retail.* Retrieved March 25, 2011, from Information Builders: http://www.informationbuilders.com/products/whitepapers/pdf/CFBI_Retail.pdf

[14] Rasmussen, N. H., Bansal, M., & Chen, C. Y. (2009). *Business Dashboards: A Visual Catalog for Design and Deployment.* John Wiley & Sons.

[15] Roberts, P. (2010). Corralling Unstructured Data for Data Warehouses. *Business Intelligence Journal*, 50-55.

[16] Scheps, S. (2008). *Business Intelligence for Dummies.* John Wiley & Sons.

Business Intelligence Systems for Strategic Management Processes: Framework and Hypotheses

Eric Y. Cheng, SUNY-Canton, 315-379-3904, cheng@canton.edu
Tai-tei Yu, Wong's International Holdings, Hong Kong, + 852-2680-3388

ABSTRACT

Business intelligence (BI) has evolved from executive information systems with wider scope and more diffused user groups. Its popularity and wide-adoption in management process has been evidenced by the tremendous growth in the supply of various BI solutions. However, its usefulness to the target users is questionable as there are still many failure cases. It is financially difficult to measure the success of BI as we were to quantify the value of decreasing uncertainties. Instead, its success is the value of its provision of timely information and relevant decision support which enable the target users at different levels across the organization to better implement their management tasks in organizational objective attainment, environmental scanning, planning, executing and controlling. In this empirical study framework, it is proposed to investigate the extent to which BI can support the management tasks in the management process. To accomplish this goal, first a thorough literature review was implemented. Second, a conceptual framework, a research model, and the hypotheses which include the factors that would affect the supports to the management tasks are presented. Finally, limitations of possible influencing factors to the success of system implementation were addressed.

Keywords: Business intelligence, strategic management, decision making, management process, executive support systems

INTRODUCTION

BI has made some enterprises more competitive, such as improved decision making, increased effectiveness, first mover advantage, leverage resources and organizational restructuring [13]. Many successful cases have demonstrated that it can improve organizational performance by speeding up decision making and/or improving the quality of management decisions in management process such as planning, environmental scanning, execution and control.

Adoption of BI is no longer the privilege of leading IT enterprises but has proliferated to medium size companies. According to a Gartner survey of more than 1,500 CIOs worldwide, BI software tops the list of technology spending priorities for companies in 2009 and that priority remains, even though IT budgets are expected to be essentially flat due to recession [6]. It is expected that the BI market to generate more than $12 billion in revenue in 2014, vs. $8.5 billion in 2008, according to

Forrester Research. Today, it appears that BI is so essential to business success that many companies cannot manage their business effectively without it [18].

LITERATURE REVIEW

Business Intelligence

A number of researchers have provided definitions for business intelligence. Raisinghani [12] defined BI as "an umbrella term that combines architectures, tools, databases, analytical tools, applications, and methodologies". Turban [15] added "It is, like DSS, a context-free expression, so it means different things to different people. Part of the confusion about BI lies in the flurry of acronyms and buzzwords that are associated with BI"

Inmon, Imhoff & Sousa [4] defined BI as "capability that help companies understand what makes the wheels of the corporation turn and help predict the future impact on current decisions". Vitt & Luckevich [16] depicted BI from three perspectives: making better decisions faster, converting data into information, and using a rational approach to management. The enablers of BI are technology, people and organizational culture.

Moss [10] described BI as an enterprise architecture for an integrated collection of operational as well as decision support applications and databases; it provides the business community easy access to their business data and allows them to make accurate business decisions. It is a new discipline, in which data is finally treated as the corporate resource. That means, any operational systems (including ERP and CRM) and any decision support applications (including data warehouses and data marts) are BI, if and only if they are developed under the umbrella and methodology of a strategic cross-organizational initiative.

Oates [11] defined BI as systems to encompass a broad range of analytical software and solutions for gathering, consolidating, analyzing and providing access to information in a way that is supposed to let enterprise's users make better business decisions. BI includes software for ETL, data warehousing, database query and reporting, multidimensional/OLAP data analysis, data mining and visualization). The key is consolidating data from many different enterprise operational systems into an enterprise data warehouse.

While most definitions are focusing on BI as a tool to transform data to information to derive knowledge through its reporting and analysis of structured information, some BI analysts took a much broader view and treat BI as including any technology, tool or application to support decision-making. Turban's definition was used for this study.

Business Intelligence and Executive Information Systems

As indicated by Turban et. al. [15], BI is the diffusion of EIS to managers at different levels with wider scope. BI has similar benefits as EIS that have been found by past researches -- but the validity of these "similar benefits" are subject to the academic proof. Most of these benefits are in fact related to providing supports for the management activities, particularly, planning, executing, evaluating and control. Singh et al [14] found that EIS was only supportive to the operational objective and strategy implementation (out of five aspects) of the strategic management process. In Singh's study, EIS was tested as to its application in the management process though it was still limited to high level strategic management. Results were interesting because EIS users did not find the benefits of EIS in all aspects when it was actually practiced in the strategic management process. This is obviously inconsistent with previous research findings -- particularly, EIS should be designed with management critical success factors in mind. These findings raised the following research questions:

- *If BI and EIS are analogous, does BI have similar benefits as EIS?*
- *BI has been adopted by users at different levels for years; do they satisfy with its performance? If yes, in which management activities?*
- *Is the huge investment in BI cost effective?*

RESEARCH MODEL and HYPOTHESES

By incorporating the BI's architecture into Hollander's classic model, a model of BI supports for management process was established (see Figure 1) -- this is the foundation of our study -- where BI plays a key role to support management process for managing the business processes and activities that are carried out for the organizational objective.

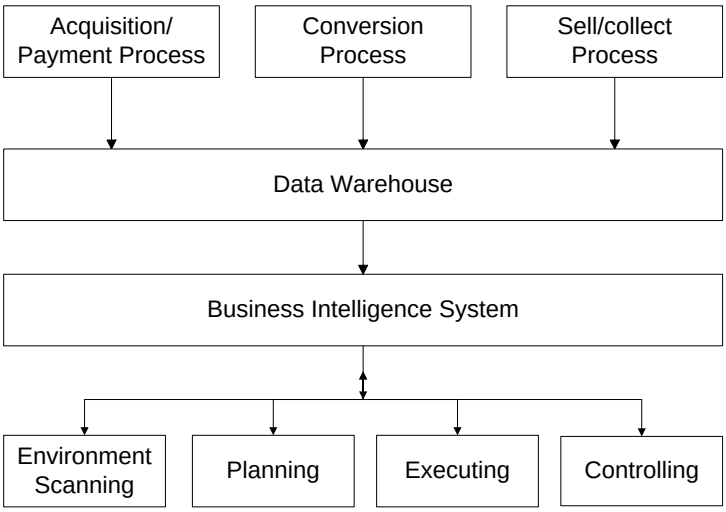


Figure 1: Integrated Model of BI Supports to Business Process and Management Functions

In the BI support model, information generated from various transaction processing systems and other information systems are stored in the data warehouse or data mart. With the BI tools (e.g. data mining, OLAP, DSS analytic), information can be extracted from the data warehouse and passed to the users for their decision-making in the management process. Navigation of information is an important feature of BI. It allows users to access to large amount of data in a quick and efficient manner [3]. Under the architecture of event-driven BI [2], information content and delivery can be more proactive, specific and focused (instead of general) to the users. Information may be: (a) operational events: being related to business opportunity/risk; (b) performance events: monitoring of deviations; (c) notification events: immediate warning of certain cases. With the event-driven capability, BI will, upon the detection of any changes, notify and alert decision-makers and keep them informed up-to-the-minute. Personalized information is pushed from the business process to the users for their timely and effective decision making no matter where they are.

Organizational objective and BI

As regards EIS's involvement in supporting the organizational objective, some EIS researchers showed that EIS does not set the objective but rather facilitates the attainment of organizational objectives through its provision of information for decision-making. According to Singh's et al study [14], there are different types of organizational objective and EIS does not necessarily support all of them. He just found the support for operational objective. If BI evolves from EIS [15], it is reasonable to assume BI to have the similar way of working as EIS. If so, we would be interested in how much (and what types of objective) it can support the objective

attainment.

Given the BI products currently available to facilitate the formulation of organizational goal in strategic planning, it is possible that BI may have such a potential capability to support the organizational objective. If users use the BI for ensuring organizational objective attainment and are satisfied with its performance in this respect, then the BI is said to be successful in this particular area. In order to prove our assumption, we devised our study with the following hypothesis.

Hypothesis 1a:	*User perceived BI success is correlated with and its supports for organizational objective attainment.*

Environmental scanning and BI

A manager acting as a decision maker in a dynamic and rapidly changing business environment should be able to scan events to understand what are the opportunities, threats, strength and weakness [5] and what actions are needed to cope with the situations [1]. In this connection, managers need a simple tool to scan all the necessary environmental information.

Given the BI technology currently available to facilitate the scanning of environmental information, it is possible that BI may have such a potential capability to support the environmental scanning. If users use the BI for the environmental scanning task and are satisfied with its performance in this respect, then the BI is said to be successful in this particular area. The following hypothesis was devised as the following.

Hypothesis 1b:	*User perceived BI success is correlated with and its supports for environmental scanning task in the management process*

Planning and BI

It appears that there is no specific direct study on the relationship between EIS and the planning process. Only Singh's et al [14] study is close to this. He could not find the supports of EIS on the strategy formulation phase of the strategic management. Senior executives did not require EIS in their policy and strategy formulation process. Or, can we say senior executives do not need EIS in their strategic planning process? Or, can we deduce that BI has the same fate as EIS? The BI industry would say no. Most prominent BI software houses are working to push out new BI products that can facilitate the planning process. If so, we would be interested in how much it can support the planning task.

Given the BI products currently available to facilitate the planning process, it is possible that BI may have such a potential capability to support the planning task. If users use the BI for the planning task and are satisfied with its performance in this respect, then the BI is said to be successful in this particular area (N.B. system success is defined and measured in Paragraph 2.4 mainly according to Lucas' [7] four measures and two-stage model). In order to prove our assumption, we devised our study with the following hypothesis.

Hypothesis 1c:	*User perceived BI success is correlated with its supports for planning task in the management process.*

Executing and BI

As to EIS's application in the executing task, EIS researchers have confirmed EIS can provide comprehensive analysis [15][17] in a personalized way [8], found causes of problems [9] in supporting executives in rolling their plans, and support the implementation of management strategy [14]. So, If BI evolves from EIS with similar functionality, there will be no reason that BI cannot support managers at different levels in executing the planned events. . If so, we would be interested in how much it can support the planning task.

Given the BI products currently available to facilitate the tracking and communicating the execution of underlying plans, it is possible that BI may have such a potential capability to support the executing task. If users use the BI for the executing task and are satisfied with its performance in this respect, then the BI is said to be successful in this particular area (N.B. system success is defined and measured in Paragraph 2.4 mainly according to Lucas' [7] four measures and two-stage model). In order to prove our assumption, we devised our study with the following hypothesis.

Hypothesis 1d:	*User perceived BI success is correlated with its supports for executing task in the management process.*

Controlling and BI

Regarding EIS's use in improving controls in organizations, researchers have confirmed its functionality like status access, trend analysis, exception reporting, tracking of critical data, drill-down capabilities, and problem explanation, etc [15]. However, Singh et al [14] did not find EIS supports in strategic control – a result inconsistent with previous papers. So, can BI support the control process? I tend to think it can because BI's capability in decision-support modeling, analytic application

and sophisticated metrics (or known as analytics/ key performance indicators) is much more powerful than EIS. So, BI has its strength in evaluating business performance strategically, tactically and operationally. In this respect, we are interested in how much it can support the controlling task.

Given quite a lot of BI products and technologies currently available to facilitate the controlling process, it is possible that BI may have such a potential capability to support the controlling task. If users use the BI for the controlling task and are satisfied with its performance in this respect, then the BI is said to be successful in this particular area (N.B. system success is defined and measured in Paragraph 2.4 mainly according to Lucas' [7] four measures and two-stage model). In order to prove our assumption, we devised our study with the following hypothesis.

Hypothesis 1e:	*User perceived BI success is correlated with its supports for controlling task in the management process*

Limitations

Researchers have identified a lot of possible influencing factors to the success of system implementation, which may also be applicable to BI system. Most of them are on behavioral issues such as:-

System decider[N1]:	System user:
- Top management support, Organizational support - Other people's opinion and influence - System decider's decision style - System decider- system researcher involvement - System decider's job characteristics - Demographics	- User's working style, personal stake - User's knowledge of system purpose - User- system researcher involvement - Organizational change caused by the system, - Other people's opinion and influence - User's job characteristics - Demographics - Organizational support

N1: System decider is defined to include system and non-system personnel who can decide on the system selection and system change.

Source: Adopted from Lucas (1990)[7]

However, our study was on the emphasis of BI system itself and human individual's characters and behaviors had nothing part of, nor consequent from, the BI system itself. Thus, we assumed all these behavioral factors constant throughout the study. This assumption is important because it can enable us to focus on the BI itself and its impact on users. Nevertheless, we have to recognize their potential impact over our results.

CONCLUSION

This research model is a good beginning at trying to understand what a successful BI should be. This study will help BI vendor, when developing a system, a tool, or a solution etc, to know what functions should be developed for which target users and what they are looking for. The users' tasks are varied in the management process. It may be planning, control, execution…etc. If implemented, variables (from the classic management theories) as the contributing factors for evaluating the success of a system in relation to the support for management process will be identified.

On the other hand, user is the key to the system success and his earlier involvement in the system development and decision stages is the added advantage. This will be proved by the change decision model and mediation model. We will be able to better understand whether user's participation system and/or professional's perception could increase the level of BI success or not.

REFERENCES

[1] Duncan, J.W., Giner, P.M. & Swayne, L.E. (1998, August). Competitive advantage and internal organizational assessment. Academy of Management Executive, 6-16

[2] Event-driven business intelligence: When timing is everything. (2001, Oct). Cognos. Available from http://www.cognos.com.

[3] Frolick, M. & Ramarapu, N.K. (1993, July). Hypermedia: The future of EIS. Journal of Systems Management.

[4] Inmon, W.H., Imhoff, C., & Sousa, R.(2001). Corporate information factory. 2nd ed., NY: Wiley.

[5] Jackson, S.E. & Dutton, J.E. (1988, September). Discerning threats and opportunities. Administrative Science Quarterly, 370-87.

[6] King, R. (2009). Business intelligence software's time is now, Bloomberg Businessweek.

[7] Lucas, H.C. Jr., Ginzberg, M.J., & Schultz, R.L. (1990). Information system implementation: Testing a structural model. NJ: Ablex Publishing Norwood.

[8] Matthews, R., & Shoebridge, A. (1992, December). EIS - A Guide for executives. Long Range Planning, 25(6), 94-101.

[9] Moynihan, G.P. (1993, July). An executive information system: Planning for post-implementation at NASA. Journal of Systems Management, 44(7), 8-14.

[10] Moss, L., & Atre, S. (2003). Business intelligence roadmap: The complete project lifecylce for decision support applications. Addison Wesley Longman.

[11] Oates, J. (2002). In Adelman, S., Moss, L., & Barbusinski, L. (2002, August). I found several definition of business intelligence

[12] Raisinghani, M. (2004). Business Intelligence in the Digital Economy. Hershey, PA: The Idea Group

[13] Rowe, J.A., Davis, S.A (1996). Intelligent information systems. CT: Greenwood Publishing.

[14] Singh, S.K., Watson, H.J. & Watson, R.T. (2001). EIS support for the strategic management process. Journal of Decision Support Systems.

[15] Turban, E., Aronson, J.E., Liang, Ting-Peng, Sharda R.(2007). Decision support and business intelligent systems. 7th edition. NJ: Prentice Hall.

[16] Vitt, E., Luckevich, M. & Misner, S. (2002). Business intelligence: Making better decisions faster. Washington: Microsoft Press Redmond.

[17] Watson, H.J., Houdeshel, G., & Rainer, K.R. (1997). Building executive information systems and other decision support application. NY: Wiley.

[18] White, C. (2002, June). Intelligent business strategies: Building the intelligent business.

A SURVEY OF OPEN SOURCE ERP SYSTEMS

Shouhong Wang
Charlton College of Business, University of Massachusetts Dartmouth
Dartmouth, MA 02747-2300 USA swang@umassd.edu

ABSTRACT

ERP systems have been widely used in business organizations. Open source ERP systems are ideal products for small and medium-sized organizations. This note provides a survey of six popular open source ERP systems on the market, and discusses our findings.

Keywords: Open source, ERP systems.

1. INTRODUCTION

ERP systems have been widely used in organizations. An ERP system integrates internal and external management information across an entire organization, embracing finance/accounting, manufacturing, sales and service, inventory management, CRM, etc. Its purpose is to facilitate the business processes of all business functions inside the boundaries of the organization and manage the connections to outside organizations. An ERP system can run on a variety of hardware and network configurations. ERP systems typically include the following characteristics.

- An ERP system is an integrated system that operates in real time.
- An ERP system has a common database which supports all applications.
- An ERP system is a set of modules with consistent appearance.
- An ERP system can be used for all types of large business organizations; however, significant work of configuration or even customization is needed to fit specific requirements for individual organizations. For instance, the configuration facility of an ERP system would allow the client organization to choose currency system, metric system, LIFO or FIFO for cost accounting, and so on. Usually, the information system development team should include specialists trained by the ERP system provider for the ERP system configuration and customization.

Recently, there are many open source ERP systems. Open source ERP systems are usually designed for small or medium-sized organizations. A business model behind open source ERP systems is the win-win relationship between the business community, the partner network, and the software editors. The partners are intended to create the market around the open source ERP system and to create services. The software editors are responsible for the quality and the vision on the development of the ERP system. The business community generates activities and contributes to the growth of the ERP system. All modules produced by the software editors, the partners and the community are to be open source. The user of an open source ERP system does

not automatically receive system support and services directly from anyone. However, an open source ERP system has its active social network that can create service offers and deals.

Selection of an open source ERP system is important to a small and midsize organization which has intention to implement an open source ERP system (Johansson and Sudzina 2008; Pobanzaou and Raymond 2011). Also, as open source ERP systems can be very useful for educational institutions to teach/learn ERP systems for many reasons such as accessibility and low cost, many MIS educators would like to see a review of open source ERP systems.

There is a lack of surveys of open source ERP systems in the literature, because such a survey is commonly not considered to be a part of academic research. There have been many surveys of open source ERP systems on the Internet. However, the reviews of open source ERP systems on the Internet are mostly posted by private companies, which could be biased. Using the Google search engine, we have found on the Internet only one relatively thorough review generated by academic scholars (Herzog 2006). However, it is at least five years old, and needs to be updated given the fast development pace of open source ERP systems as well as the IT environment.

This note provides a survey of six popular open source ERP systems. As Microsoft Dynamics is a major ERP system for small and medium-sized organizations and is a strong competitor of these open source ERP systems, Microsoft Dynamics is overviewed briefly. After the review of the six popular open source ERP systems, overall findings are discussed.

2. SAMPLING AND COMPARISON CRITERIA

The selection of samples of open source ERP systems for this survey went through two stages. In the first stage, general searches on the Internet were conducted to find 25 open source ERP systems that were well recognized by the industry (e.g., on the Wikipedia). In the second stage, each of these open source ERP systems was examined in great detail based on its demo. If an open source ERP system did not provide informative demo on the Internet or its demo was obviously inferior, it was not selected for further comparison. As a result, six open source ERP systems were used in this comparison.

Considering the current development of the IT environment, the trends of social networks, and the converging ERP systems functionalities, we applied the following four criteria in comparing these six open source ERP systems.

(1) Readiness for cloud computing being a Web-based system.
(2) Completeness of commonly required ERP functionalities.
(3) Easiness of configuration on any operating systems.
(4) Active and large social network for users community support.

3. REVIEW OF OPEN SOURCE ERP SYSTEMS

In this section, we use many screenshots for visual illustration. As discussed at the end of this

note, webERP is considered to be the best one based on our review. Thus, common ERP functionalities are illustrated by the screenshots of webERP. For each of other open source ERP systems, only screenshots with unique features are displayed.

3.1. Microsoft Dynamics as a strong competitor of open source ERP systems

Microsoft Dynamics (http://www.microsoft.com/en-us/dynamics/default.aspx) provides affordable scalability for small and midsize business. With benchmarked performance up to 2,250 users, Microsoft Dynamics provide a variety of functionality across financials, supply chain management, and customer relationship management. Its database management system is Microsoft SQL Server.

The ownership costs for Microsoft Dynamics applications are much lower than other ERP systems such as SAP and Oracle. Microsoft Dynamics certification program offers solutions designed to meet the needs of specific industries.

As Microsoft Office and Windows have already widely been used in many organizations if not all, Microsoft Dynamics seems to be easy for system integration for these organizations. Microsoft is considered to be one the most reliable software provider in the world.

3.2. webERP

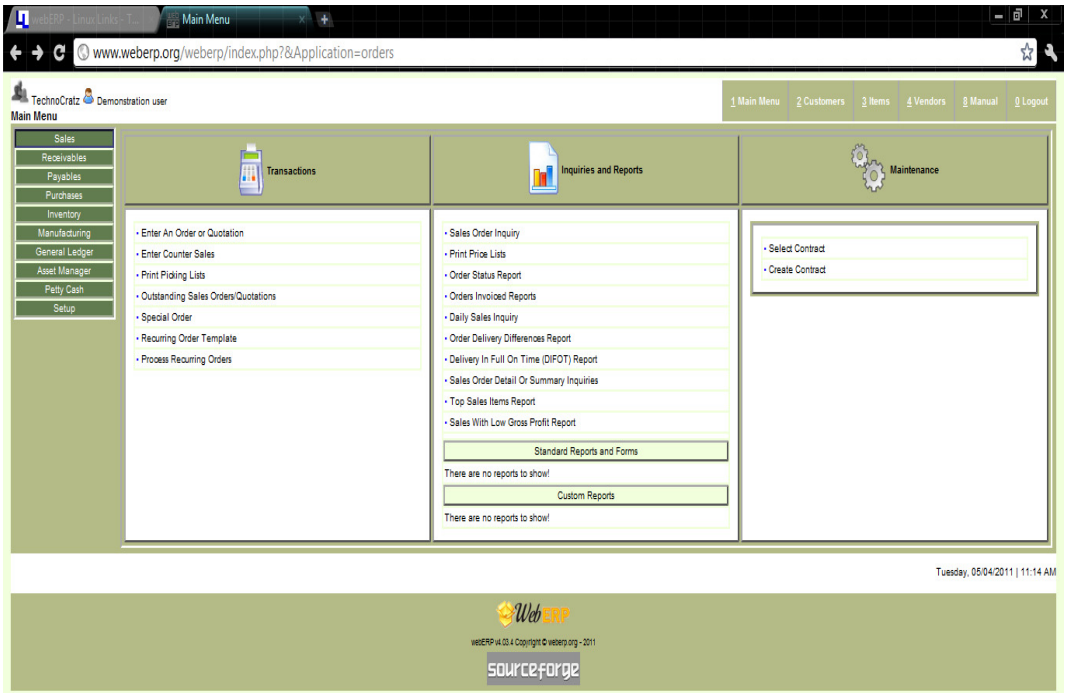
webERP (http://www.weberp.org/) is a Web-based open source ERP system. It is implemented by using PHP and MySQL. There have been more than 128,000 downloads of the system to date. It supports multiple languages. Besides the official Web site of webERP, there are two main types of support of webERP.

(1) Community Support - There are mailing lists that consist of all the developers and users of webERP. The webERP mailing lists can be accessed via the nabble forum (www.nabble.com), and are the best places to inquire about any issue you may have with using or developing webERP. The users archives and developer archives of the mailing lists contain valuable knowledge.

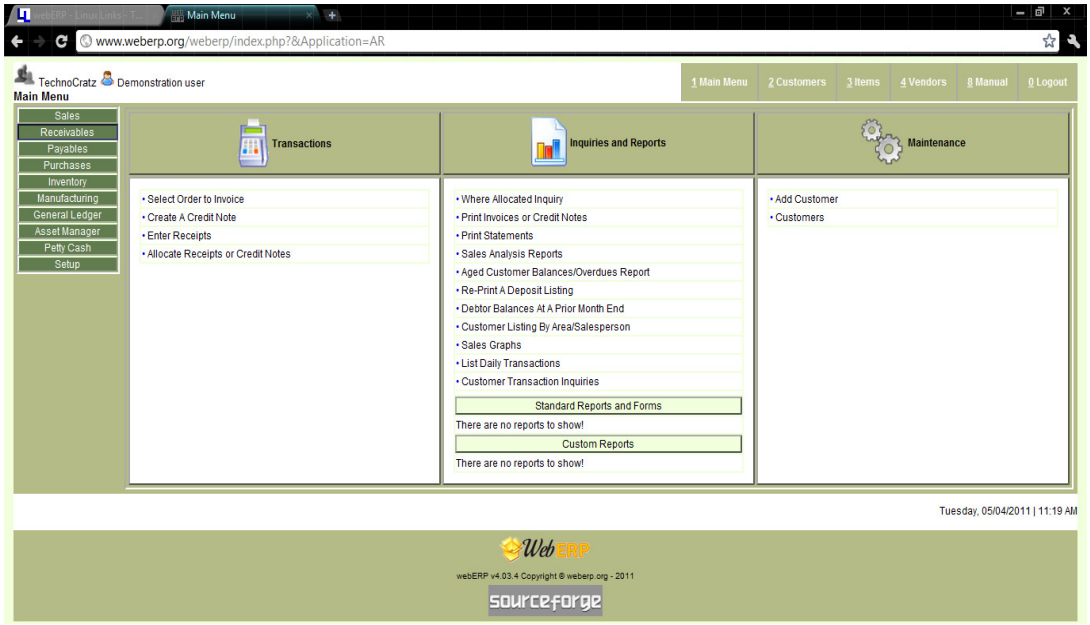
(2) Commercial Support - Several companies offer commercial support which may be preferable for ordinary business organizations.

webERP is easy to operate. The Main Menu has a standard display of three categories: Transactions, Inquiries, and Reports and Maintenance.

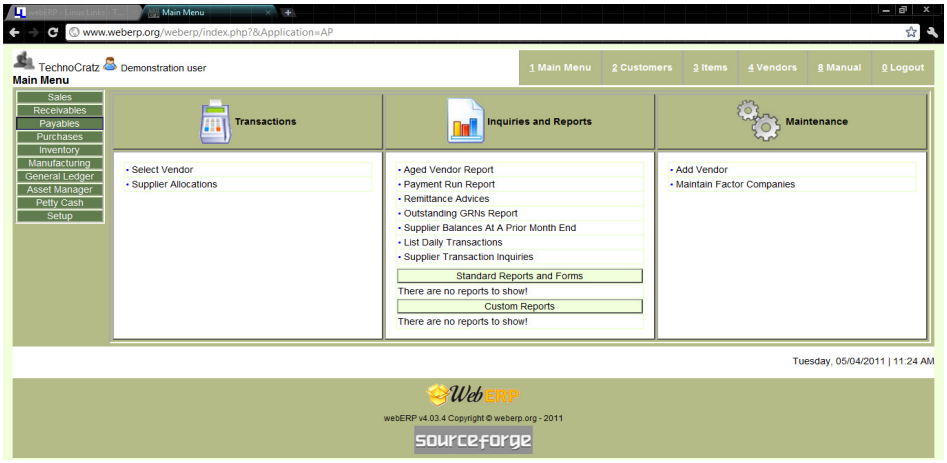
Sales dashboard – It lists orders, quotations, and summaries of sales orders, invoices, prices, and other report links.



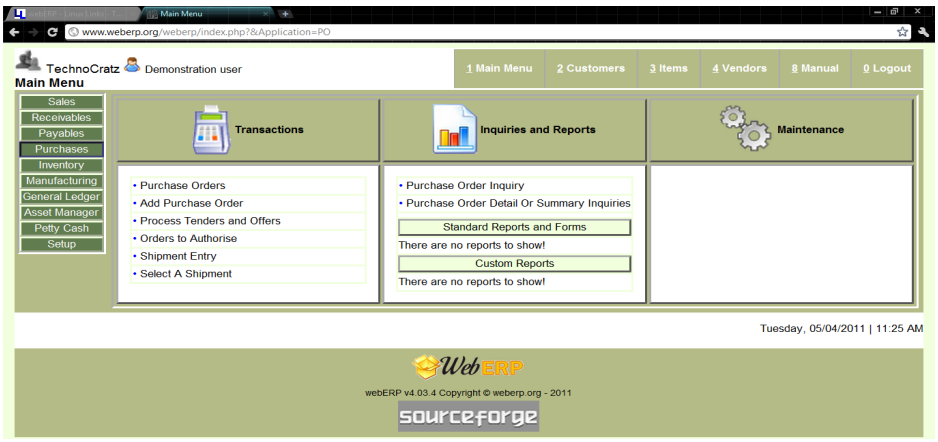
Receivables dashboard – It links to receipts, invoices, statements, overdue records, daily transactions, and others.



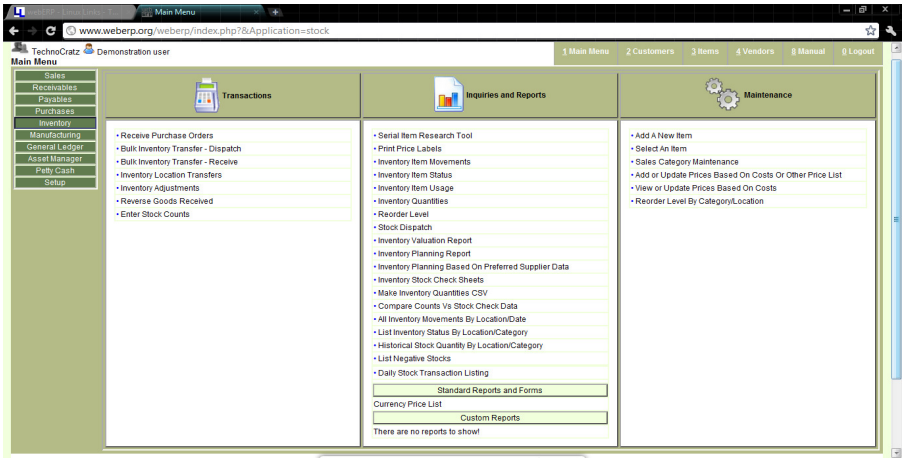
Payables dashboard - It contains vendor related links similar to receivables dashboard.



Purchases dashboard - It shows the status of various purchase orders and shipment entries.



Inventory dashboard – It shows inventory related transactions and reports. It also allows the user to add, update, or delete inventory items.



Manufacturing dashboard - It links to work order entry, materials inquiries, and other reports. It allows the user to perform MRP calculation.

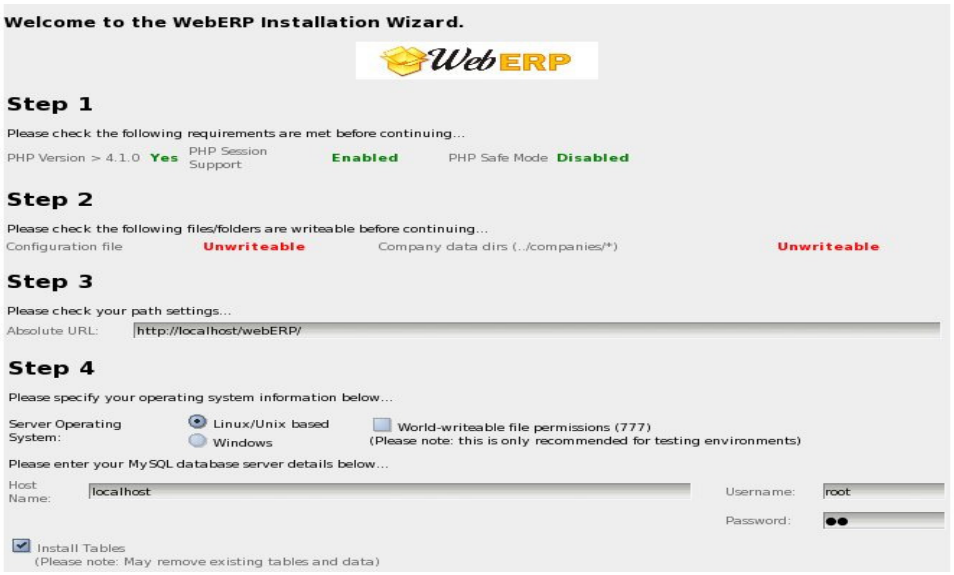
General ledger dashboard - It links to all entries of customer and supplier invoices, refunds, payment, bank statements, account inquiries, and others.

Asset manager dashboard - It has asset related links that allow the user to add or delete asset, change asset location, and other operations.

Before an ERP system can operate, configuring must be done to set many attributes or parameters (e.g., structures, terminology, access rights, etc.) of the system for the organization to meet the specific needs in the best possible way. webERP has a central dashboard for links of configuring functions. The following screenshots are examples of configuring webERP.

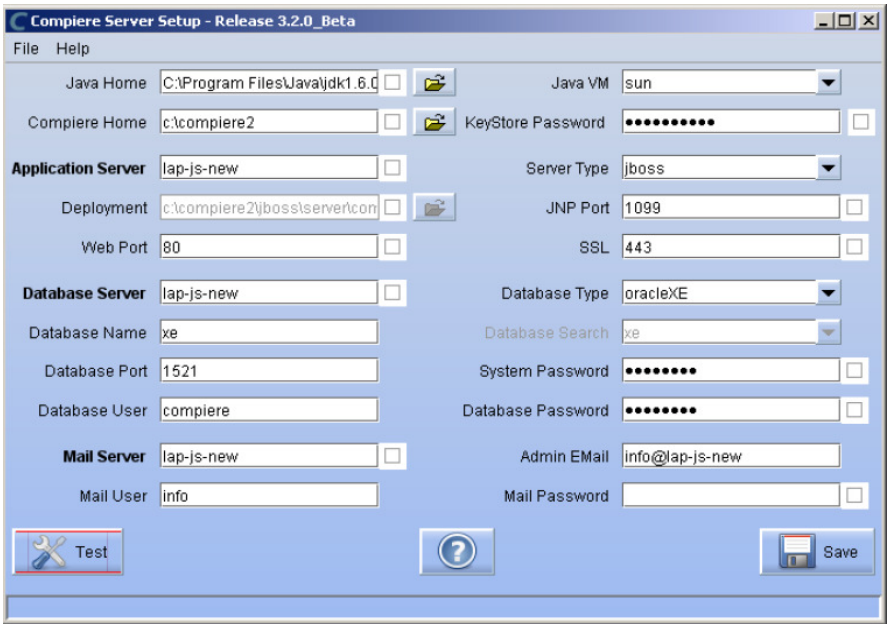
Set up dashboard – It links to configure operations for entire ERP settings.

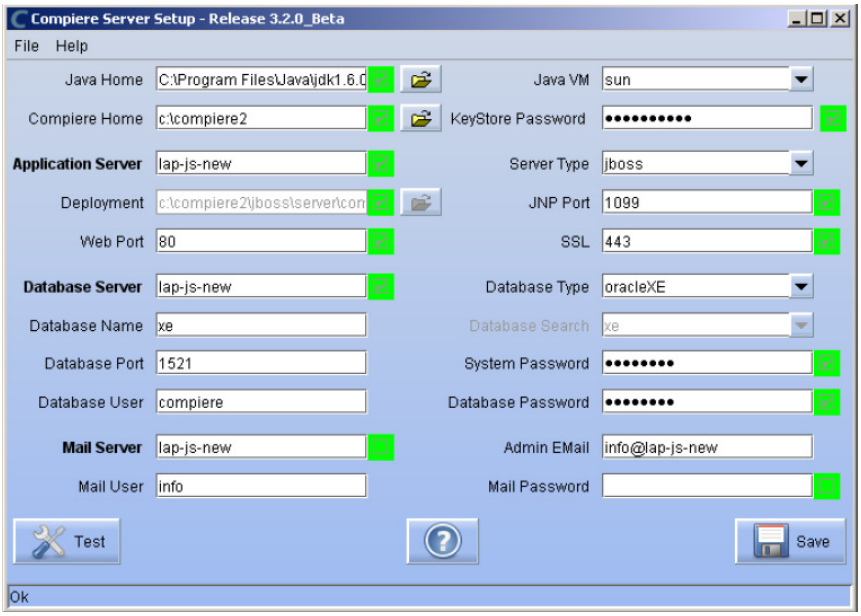
Configuring **-** The use of Installation Wizard

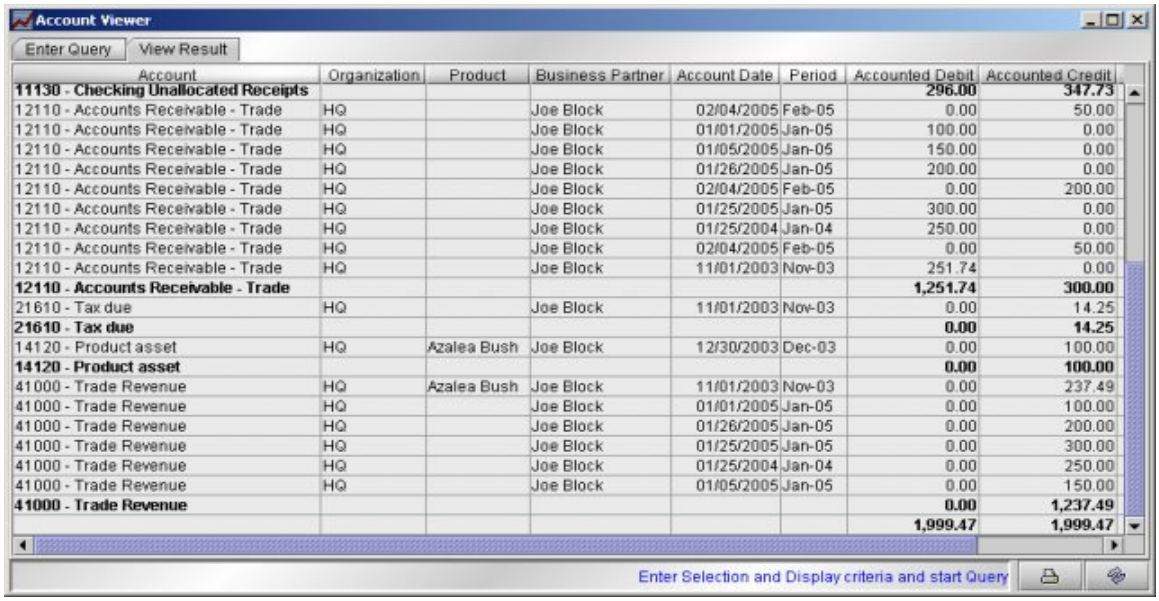


3.3. Compiere

Compiere (http://www.compiere.com) is a Java based system. There have been over 1.35 million downloads of Compiere. Compiere is a Web-based open source ERP system. There is a free community version without support as well as commercial editions that provide complete support and guarantee, training resources, and other services. The following is several samples of the user interface.



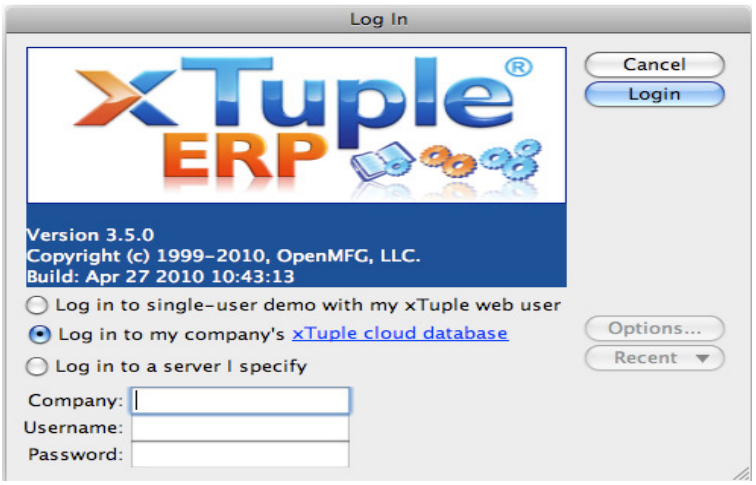


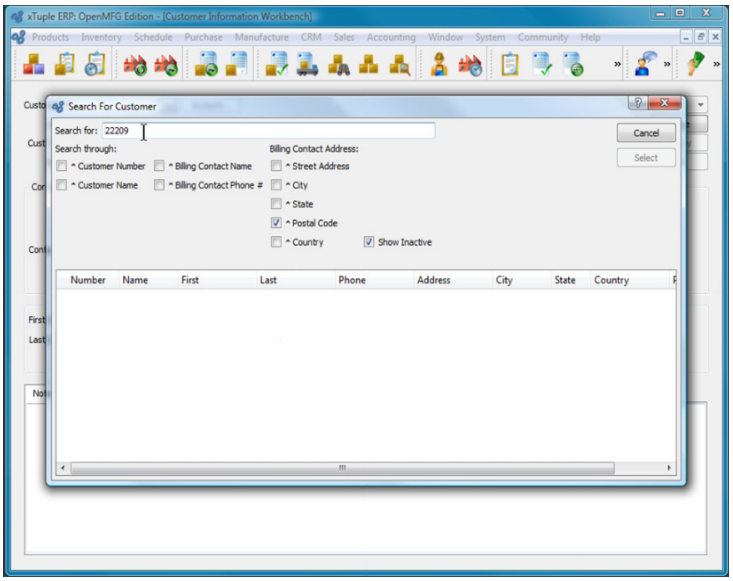


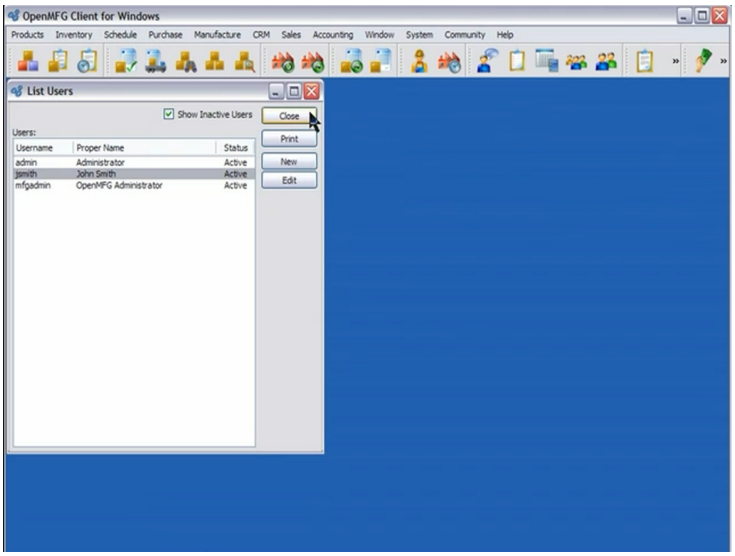
Account	Organization	Product	Business Partner	Account Date	Period	Accounted Debit	Accounted Credit
11130 - Checking Unallocated Receipts						**296.00**	**347.73**
12110 - Accounts Receivable - Trade	HQ		Joe Block	02/04/2005	Feb-05	0.00	50.00
12110 - Accounts Receivable - Trade	HQ		Joe Block	01/01/2005	Jan-05	100.00	0.00
12110 - Accounts Receivable - Trade	HQ		Joe Block	01/05/2005	Jan-05	150.00	0.00
12110 - Accounts Receivable - Trade	HQ		Joe Block	01/26/2005	Jan-05	200.00	0.00
12110 - Accounts Receivable - Trade	HQ		Joe Block	02/04/2005	Feb-05	0.00	200.00
12110 - Accounts Receivable - Trade	HQ		Joe Block	01/25/2005	Jan-05	300.00	0.00
12110 - Accounts Receivable - Trade	HQ		Joe Block	01/25/2004	Jan-04	250.00	0.00
12110 - Accounts Receivable - Trade	HQ		Joe Block	02/04/2005	Feb-05	0.00	50.00
12110 - Accounts Receivable - Trade	HQ		Joe Block	11/01/2003	Nov-03	251.74	0.00
12110 - Accounts Receivable - Trade						**1,251.74**	**300.00**
21610 - Tax due	HQ		Joe Block	11/01/2003	Nov-03	0.00	14.25
21610 - Tax due						**0.00**	**14.25**
14120 - Product asset	HQ	Azalea Bush	Joe Block	12/30/2003	Dec-03	0.00	100.00
14120 - Product asset						**0.00**	**100.00**
41000 - Trade Revenue	HQ	Azalea Bush	Joe Block	11/01/2003	Nov-03	0.00	237.49
41000 - Trade Revenue	HQ		Joe Block	01/01/2005	Jan-05	0.00	100.00
41000 - Trade Revenue	HQ		Joe Block	01/26/2005	Jan-05	0.00	200.00
41000 - Trade Revenue	HQ		Joe Block	01/25/2005	Jan-05	0.00	300.00
41000 - Trade Revenue	HQ		Joe Block	01/25/2004	Jan-04	0.00	250.00
41000 - Trade Revenue	HQ		Joe Block	01/05/2005	Jan-05	0.00	150.00
41000 - Trade Revenue						**0.00**	**1,237.49**
						1,999.47	**1,999.47**

3.4. PostBooks

PostBooks (http://www.xtuple.com/) is based on the award winning xTuple ERP Suite. Its foundation is open source PostgreSQL database and the open source Qt framework for C++. It supports multiple languages. Its ERP functions are quite typical. The following is several samples of the user interface.

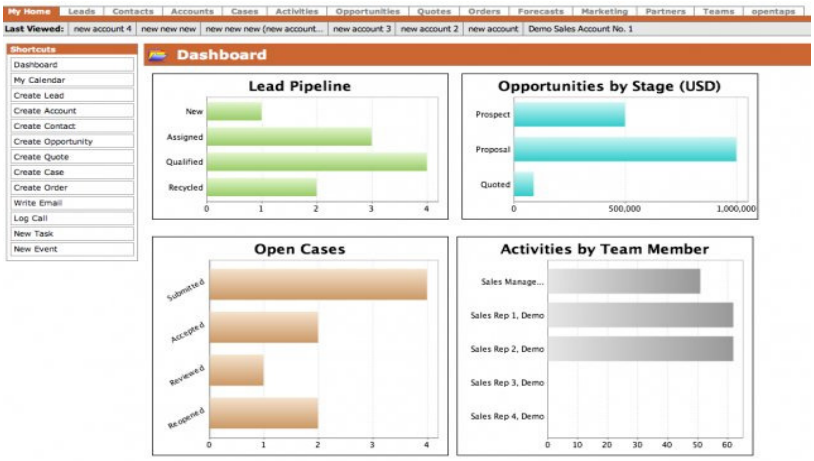

Log In
Cancel
Login
Version 3.5.0
Copyright (c) 1999-2010, OpenMFG, LLC.
Build: Apr 27 2010 10:43:13
Log in to single-user demo with my xTuple web user
Log in to my company's xTuple cloud database
Log in to a server I specify
Options...
Recent
Company:
Username:
Password:

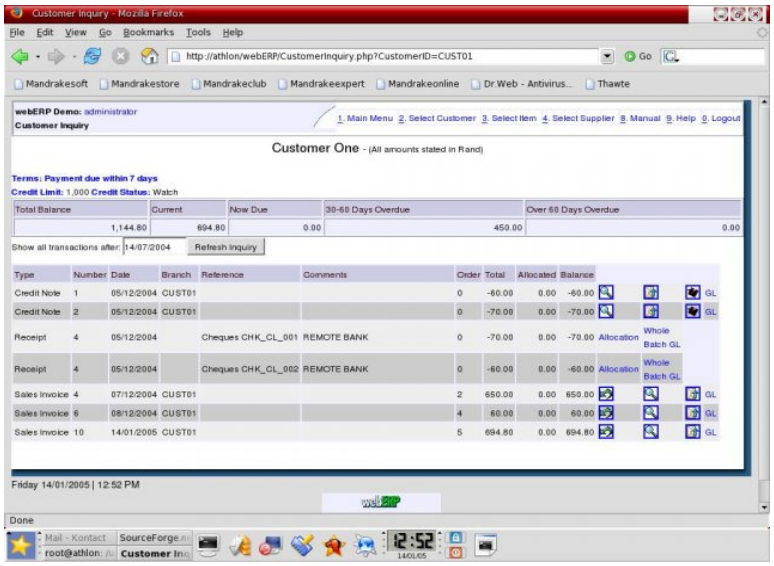

Search For Customer
Search for: 22209
Search through:
Billing Contact Address:
Cancel
Select
Show Inactive
Number
Name
First
Last
Phone
Address
City
State
Country


OpenMFG Client for Windows
List Users
Show Inactive Users
Close
Print
New
Edit
Users:
Username
Proper Name
Status
admin
Administrator
Active
jsmith
John Smith
Active
mfgadmin
OpenMFG Administrator
Active

3.5. Opentaps

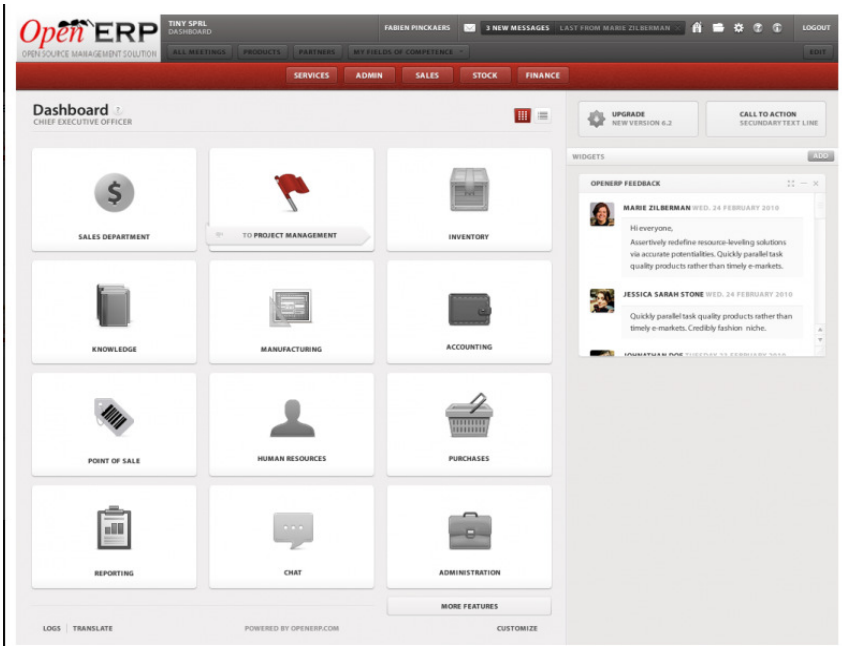
Opentaps (http://shop.opentaps.org/) incorporates several open source projects, including Apache Geronimo, Tomcat, and OFBiz for the data model and transaction framework; Pentaho and JasperReports for business intelligence; Funambol for mobile device and Outlook integration. Opentaps applications which provide user-driven applications for CRM, accounting and finance, warehouse and manufacturing, and purchasing and supply chain management. Interestingly, it has Asterisk PBX integration. The following is several samples of the user interface.

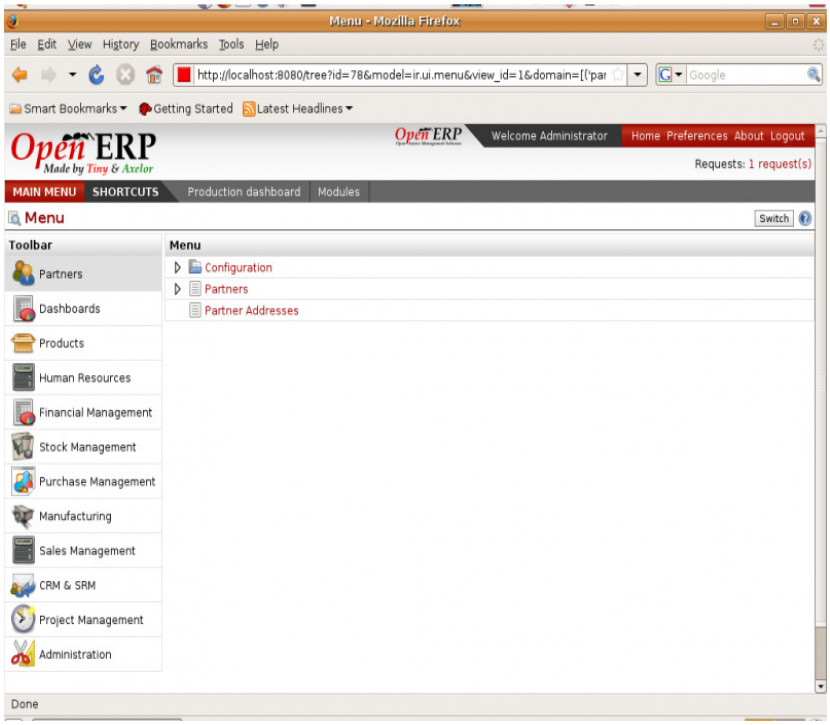


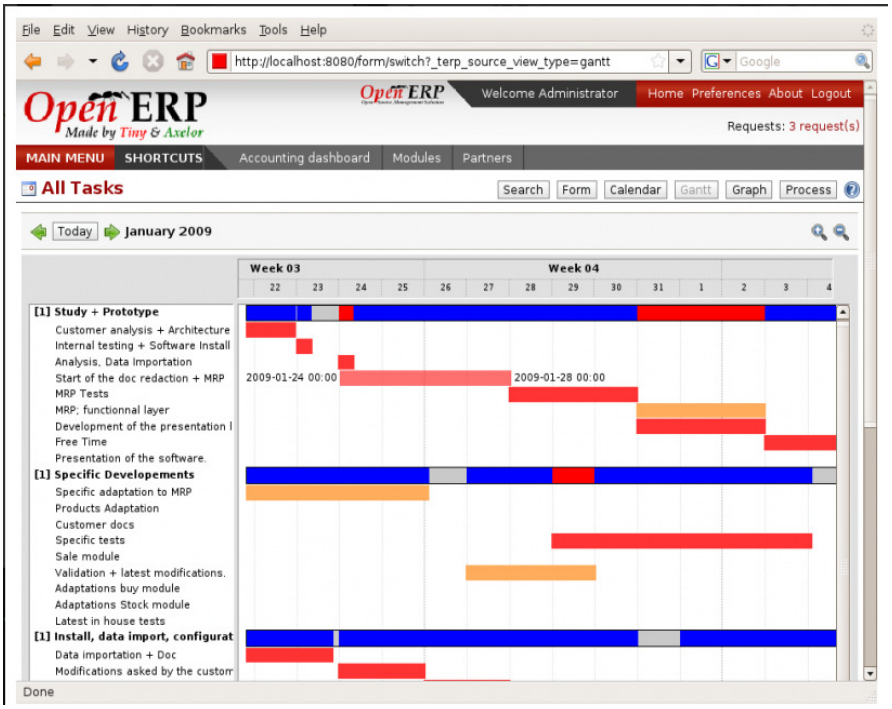


3.6. OpenERP

OpenERP (http://www.openerp.com/) is open source ERP suit that supports enterprise modules, logistics, accounting and finance, HRM, CRM, project management, and other business processes. It is complete package with commercial version available.

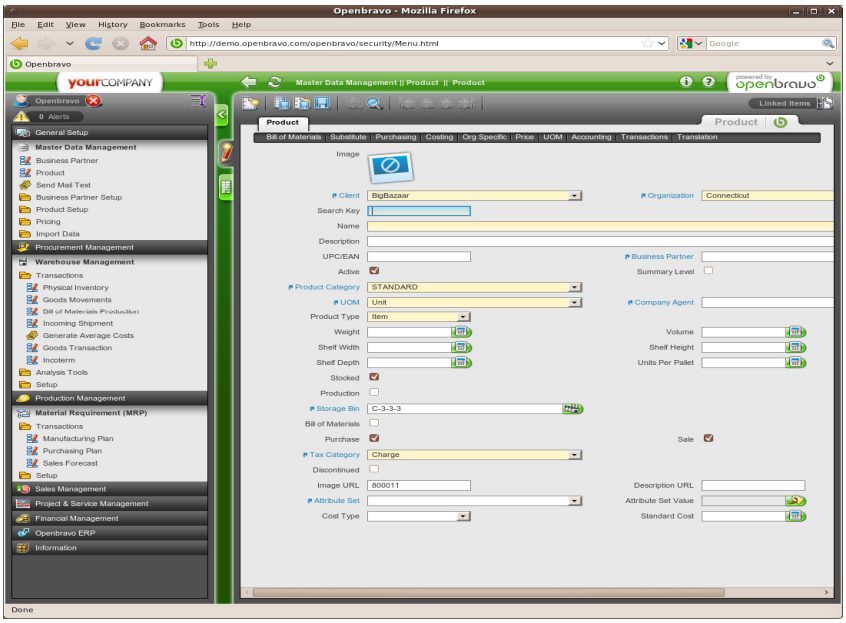

Dashboard
SALES DEPARTMENT
TO PROJECT MANAGEMENT
INVENTORY
KNOWLEDGE
MANUFACTURING
ACCOUNTING
POINT OF SALE
HUMAN RESOURCES
PURCHASES
REPORTING
CHAT
ADMINISTRATION
MORE FEATURES


Menu - Mozilla Firefox
File Edit View History Bookmarks Tools Help
Made by Tiny & Axelor
Welcome Administrator
Home Preferences About Logout
Requests: 1 request(s)
MAIN MENU
SHORTCUTS
Production dashboard
Modules
Menu
Switch
Toolbar
Partners
Dashboards
Products
Human Resources
Financial Management
Stock Management
Purchase Management
Manufacturing
Sales Management
CRM & SRM
Project Management
Administration
Configuration
Partners
Partner Addresses
Done



3.7. OpenBravo

OpenBravo (http://www.openbravo.com/) is a web based ERP solution. It originally was based on Compiere. It supports standard ERP features like production information, inventory, customer information, order tracking, and workflow information.



4. DISCUSSION

Obviously, the major advantage of open source ERP systems over commercial ERP systems is free to use. However, the disadvantages of open source ERP systems are also apparent, such as lack of support and lack of active maintenance. While providing common ERP functionalities, each open source ERP system has its unique characteristics. Nevertheless, based on our review of several popular open source ERP systems with various features, webERP seems to the best among others in the four critical aspects: web-based and readiness for cloud computing, ERP functionalities, easiness of configuration, and users' community network.

(1) webERP is a pioneer of cloud computing where the application can take place anywhere on the Internet and be delivered to the business through "the cloud". It requires only a Web browser and PDF reader to use. The system is easily available online and even the download size is comparatively less and other system requirements.

(2) webERP provides common ERP systems functionalities, with emphasis on accounting, support various business processes of wholesale, distribution, and manufacturing in order entry, taxation, general ledger, accounts payable/receivable, inventory, purchasing, banking, production, contract costing, fixed assets management, etc.

(3) webERP can be easily configured on any operating systems and the computer processing requirements are light. webERP can be run over an internal LAN. It can also be run on a third party web-hosting server external to the business which needs only a router and connection to the Internet to use webERP.

(4) webERP has built up a good network for users community support. Several software companies offer commercial support for webERP.

The future work of this study includes a stronger literature review and rationale for considering the factors to evaluate the different open source ERP systems, and detailed ranks the open source ERP systems based on the survey.

Acknowledgement: Mr. Sreevatsan Narayanan has made valuable contributions to this survey study.

REFERENCES*

[1] Compiere: <http://www.compiere.com>.
[2] Herzog, T., A Comparison of Open Source ERP Systems, June 2006. <citeseerx.ist.psu.edu/viewdoc/download?>.
[3] Johansson, B. and Sudzina, F., ERP systems and open source: an initial review and some implications for SMEs, *Journal of Enterprise Information Management*, 21(6), 649-658.
[4] Microsoft Dynamics: <http://www.microsoft.com/en-us/dynamics/default.aspx>.

[5] OpenBravo: <http://www.openbravo.com/>.
[6] OpenERP: <http://www.openerp.com/>.
[7] Opentaps: <http://shop.opentaps.org/>.
[8] Poba-nzaou, P. and Raymond, L., Managing ERP system risk in SMEs: a multiple case study, *Journal of Information Technology*, 26(3), 2011, 170-192.
[9] PostBooks: <http://www.xtuple.com/>.
[10] webERP: <http://www.weberp.org/>.
[11] Wikipedia: List of ERP software packages, <http://en.wikipedia.org/wiki/>

(* All Web sites cited are accessed in October 2011)

IMPROVING THE BALANCE: ACTIONS TO SUPPORT AND SUSTAIN WOMEN IN UNDER-REPRESENTED FIELDS

So Phi A Ngo, RWU '12, Roger Williams University, sngo940@g.rwu.edu
Jessica Peacock, RWU '12, Roger Williams University, jpeacock144@g.rwu.edu
W. Brett McKenzie, wmkenzie@rwu.edu (401) 254-3534
Gabelli School of Business, Roger Williams University,
One Old Ferry Road, Bristol RI 02809

ABSTRACT

Women in the computing field have organized to systematically address the underrepresentation of women in technical careers. To this end they have created the industry and academic partnerships described in this paper. As a comparison to computing, the authors look at the field of finance, both from an industry and academic perspective. Drawing on personal experience and peer interviews, the undergraduates who will present this paper, will contrast their experiences in computing with those of their peers in the finance industry.

INTRODUCTION

Despite dramatic changes in work and career choices for women since the second wave of feminism in the nineteen seventies, there continue to be areas where women remain under-represented. In some careers, women within a field have banded together to highlight the issue and to provide support for women beginning their careers. Computing is one of the fields where the presence of women has been in decline since the nineteen eighties and career women have tried to address the issues.

Background

The CRA Taulbee Study [1], is the definitive compilation of female participation in computing areas. The most recent data shows that 11.8% of bachelor's degrees in the computing and information sciences are being awarded to women. This is a decline of almost 50% in the last ten years. Similar declines are seen in the awarding of associate's and masters' degrees. The most recent data, compared with NSF STEM data, shows that the increase in women seeking science and engineering degrees is not matched by women seeking computing degrees.

As a consequence of the underrepresentation of women in computing, various organizations, many funded by the National Science Foundation, have been initiated to determine the causes for women to reject computing as a career choice and to discover ways to promote computing as a career for women. Leaders in the field include the Anita Borg Institute, founded in 1997 [2] and the National Center for Women in Information Technology founded in 2004 [3]. These organizations combine forces in the Grace Hopper Celebration of Women in Computing (GHC) [4], a national conference that focuses on the successes of women and the pathways that women can follow to compete and excel in the computing field. While the complete causes of the

avoidance of computing by women is still not agreed upon, it appears that unconscious bias, from the promoting of computing for boys to the plethora of casual computing in games designed for males, has some bearing on the current state.

Similar to computing, finance has been noted as being a bastion of male employment, especially in popular culture, with films such as *Wall Street* (1987, 2010) or books such as *Tales from the Boom Boom Room* (2002). In the recent downturn, women have lost positions at a surprising rate and, more importantly, are not recovering at the same rate [5]. Comparing computing with finance is challenging, especially because finance is mature, for example the insurance industry traces itself to at least 1688 with the founding of Lloyds of London, and because its components, such as insurance, banking, or investments, are more distinct than segments of computing.

There are organizations, similar to those mentioned in computing, that have developed to support women in finance. The Financial Women's Association [6] founded in 1956 and the Forte Foundation [7] founded in 2001 are among the leaders. Similar to GHC, the financial industry in New York hosts the WOW (Women of Wall Street) conference. The public/private collaboration with academic and industry partners is not as prominent among finance as it is in computing. The focus of the finance organizations, too, is more on individual advancement than industry or social change.

RESEARCH METHODS AND PROCEDURES

Methods

The primary research method is personal ethnography. The principals, due to gender and interest in computing, have been able to participate in some of the support events for women considering computing. One of the researchers has a particular advantage as an international student in that she is able to contrast American experiences with expectations from another culture.

Data Collection

Primary data collection derives from participation in the events to support women in computing including international conferences and targeted workshops. Supportive data collection is from interviews with colleagues participating in the same or similar workshops.

Secondary data collection, mostly for the finance section, is from advertised activities and reports on those activities. Interview with peer females students in the finance field provides for informants in the comparative field.

REFERENCES

[1] Vegso, J. (2008) CRA Taulbee Trends: Female Students & Faculty http://www.cra.org/resources/taulbee/

[2] Anita Borg Institute for Women in Technology, http://anitaborg.org/about/history/

[3] National Center for Women in Information Technology http://www.ncwit.com/about.factsheet.html

[4] Grace Hopper Conference (2012), http://gracehopper.org/2012/

[5] Jacobs, D. L. (2011) " 'Mancession' Fades As More Men Than Women Find Jobs" http://www.forbes.com/sites/deborahljacobs/2011/12/06/economic-recovery-is-gender-biased-study-suggests/

[6] Financial Women's Association, http://www.fwa.org/aboutus/aboutus.htm

[7] Forte Foundation, http://www.fortefoundation.org/site/PageServer?pagename=abt_forte

A METHODOLOGY FOR DEVELOPING AND ASSESSING CRITICAL THINKING SKILLS IN BUSINESS STUDENTS USING SUBSTANTIVE COURSE MATERIAL

Martha S. Weisel
Associate Professor of Legal Studies in Business
Hofstra University
Hempstead, New York 11550
516-463-5666
actmsw@hofstra.edu

ABSTRACT

Critical thinking skills are an important component of higher level learning. Students need to be able to evaluate complex situations in their professional and personal lives. Using the substantive material in a required introductory business course, the author developed a model that concentrated on developing specific critical thinking skills. This system was used by the students in weekly homework assignments based on the course material taught during the week. The problems were graded after the semester using a rubric developed by the author. At the conclusion of the study, it appeared that the teaching technique had some success in developing specific critical thinking skills, but that further study was needed.

Critical thinking substantive course material rubric business students

INTRODUCTION

Criticism of American education and concern for the ability of our graduates to compete successfully in a global environment is not a new story. In the 1950's the launch of Sputnik shook United States education as it highlighted American students' failure to keep up with their Soviet counterparts. [6] Fast forward to the present as American students are being portrayed as lagging behind their Chinese counterparts. [15] Newscasters, academics, politicians and ordinary citizens have addressed the issue. Increasingly, business leaders have expressed concerns about the education being received by business students. Some of the concerns have focused on poor problem solving and strategic planning skills, arguing that the world economy places great emphasis on critical thinking and this need is not being met. This is based in part on an understanding that success in the business world requires students to exercise critical thinking skills in their decision making. Some argue that analysis, skepticism and judgment are as important if not more important than functional knowledge. [10,p.58]

Association to Advance Collegiate Schools of Business (AACSB)

One of the institutions that focuses on the need to maintain high quality business education and improve the abilities of business school graduates is the AACSB, the organization that accredits schools of business in the United States as well as international business schools. [12, p.113] AACSB Standard 15 emphasizes the requirement that business schools teach students to develop critical thinking skills as part of the undergraduate curriculum. In determining that critical thinking is an essential learning objective, the AACSB acknowledged that due to the complexities involved in developing critical thinking skills, different universities use different approaches and emphasize different skills in achieving the desired goals. [1, p.62] Accredited

schools must provide outcome assessments showing the development of critical thinking skills. [8, p. 313] Despite the importance of development of critical thinking skills, the AACSB provides no specific definition of critical thinking and no measures as to evaluate and assess whether or not students have achieved this learning objective.

DEFINING CRITICAL THINKING

There are concepts that we cannot define, but that we recognize when we see them. The literature is replete with various definitions for critical thinking, but no definition truly explains all of the various aspects of critical thinking. Some definitions focus on specific skills, such as problem solving, others define critical thinking as creating a mindset that looks at information in a questioning manner. Volumes have been written by scholars as to the various elements that make up critical thinking skills. One of the most respected is Taxonomy of Educational Objectives. [2] The Taxonomy formulated a six-part process by which students are taught to learn. These six skills have been standardized to include knowledge, comprehension, application, analysis, synthesis and evaluation. At its most basic, students should be able to memorize facts and be able to repeat them. Comprehension allows students to understand the facts that they have been taught and application gives students the ability apply those facts to different situations (fact based learning). Critical thinking skills take students beyond memorization, comprehension and application of facts. Students must be able to analyze certain situations using the information they need to know. By analyzing a situation, a student should be able to put all of the information together and sort it out (synthesis) so that a conclusion can be reached.

Teaching Critical Thinking Skills

The question raised is how can business students develop and refine their critical thinking skills at the undergraduate level as part of a business curriculum. Often, students are not directly taught critical thinking as the focus in most classes is on the development of knowledge. However, students are expected to display critical thinking skills in at least some of their assignments. [11,p.172] In an effort to deal with this situation, many professors have made efforts to specifically teach critical thinking skills within the framework of a specific course. A finance class used client memos to assess critical thinking skills of finance majors. [5, p.14] Debates have been used in several studies as a tool for developing critical thinking skills. [14, p.77][4, p.658] The Wall Street Journal has been used in a business communication class. [9, p.67]

Assessment

Although a wide variety of approaches have been used in business classes as a way of teaching critical thinking, few have been assessed objectively and there is little evidence of improvement. [3, p.235] Available tests do not measure the ways in which students arrive at their answers which may be one reason why there are few empirical studies on critical thinking in business education. [11, p.3] Some of the projects had no assessment mechanism. Other studies have used anecdotal reports and self evaluations. [4, p. 658]. Rubrics were used in a capstone course developed to assess critical thinking but the authors concluded that the rubrics are difficult to use and need to be modified frequently, making critical thinking difficult to assess. [8, p. 317] Finding ways to evaluate the development of critical thinking skills in business education may be the most difficult task in planning and implementing research. Critical thinking involves so

many components that the first thing that must be considered is the specific objective that will taught and assessed. [7, p. 300] Activities need to be developed that measure a broad range of tasks representing diverse subject matter and different skills. The assessment tool must be able to interpret the merit of different responses. [11,p.177]

METHODOLOGY

The research indicates that although students are expected to display critical thinking skills, most students are not directly taught these skills in business classes as the focus tends to be on the development of knowledge. This study was designed to determine if specific critical thinking skills can be successfully taught within the substantive curriculum of a business course. The study focused on the ability to make inferences, a specific skill targeted in certain critical thinking tests (Watson-Glaser). In order to analyze a situation, one must be able to make inferences based on specific information. Students need to determine what information is needed to reach a well thought out decision. Students were encouraged to determine information that was unclear (ambiguous) as well as information that was missing.

Design of the problem

A fact pattern was created for each chapter of the required text so that the substantive material in the class was used to develop critical thinking skills. Rather than using a narrative, students were given a specific rubric to be used in answering the questions. The rubric was provided via Blackboard, as follows:

1. Facts given – Students were required to state 2 facts that they thought were significant
2. Missing information needed to reach a conclusion
3. Information that needs to be clarified before reaching a conclusion
4. My conclusion

In reaching his/her conclusion, each student was required to explain the way in which missing information and/or ambiguous information was used to form a decision.

During the first day of class one fact pattern was reviewed in class. Emphasis was placed on reaching a decision based on something more than personal opinion. What additional information would help you reach a decision? What about the fact pattern was unclear? Students were then given the rubric that was to be used in answering each fact pattern and were assigned that fact pattern as homework for the next class. During the next class, there was a discussion as to ways in which conclusions were reached. Did the conclusion evaluate additional information? Did clarification of ambiguities prove helpful in reaching a conclusion?

After the first week of class, students were required to submit one assignment each week using the four part rubric. The fact pattern was based on a substantive topic that was taught during the week. The assignments were collected and counted as homework. They were not graded, the only requirement was to follow the rubric. Prior to the return of each assignment, ten to fifteen minutes of class time was used to discuss their decisions and the basis on which their decisions were reached. Each fact pattern was based on an actual case.

Collection of data

Copies were made of the students work so that grading could be done after the semester was completed. To grade each assignment, I wrote up answers for each part of the question, listing

facts, missing information, unclear information and possible conclusions. My teaching assistant did the same thing independently. It should be noted that my teaching assistant had been in my class as a freshman and was familiar with many of the cases. Both sets of answers were put together before grading each assignment. Each assignment was graded by me and then independently by my teaching assistant. Four assignments were graded, the first assignment, after Week four, after Week eight and the last assignment.

ANALYSIS OF THE DATA

The practice assignment was done in class and independently. Most students were able to identify only one fact. Over 50% of the students were unable to determine any missing or ambiguous information. As a result, over 50% of the students were unable to use any information in reaching a conclusion. Please note the differences in my grading (faculty) and my teaching assistant. Although we both used the identical guidelines, there is a substantial difference in grading. For example, I determined that 18 students were unable to reach a conclusion using any missing/ambiguous information. My teaching assistant determined that 10 students were unable to reach a conclusion using the missing/ambiguous information.

Practice assignment – Day 1

Topic: Smoking (Role of the United States Supreme Court)

FACULTY

No. Correct	Facts	Missing Info	Ambiguous Info	Conclusion
0	2	19	14	18
1	25	8	14	10
2	1	0	0	0

STUDENT

No. Correct	Facts	Missing Info	Ambiguous Info	Conclusion
0	2	19	12	7
1	25	8	10	19
2	1	0	2	1

Assignment – Week 1

The first assignment shows that most students were able to determine the facts in the assigned problem. Over 50% of the students were able to identify two relevant facts. This was a significant improvement. Only 5 students were unable to identify any facts. However, it should be noted that the differences in grading between faculty/ TA persist, even as to the facts. On assignment 1 students were unable to determine what missing and/or ambiguous information was needed to reach a decision. The grades for assignment 1 were lower than the grades on the practice assignment. This is not a surprise. The practice assignment was done in class first so that students who took notes were able to repeat the information on the assignment that was handed in during the next class. This option was not available for Assignment 1.

Topic: Workplace Safety (Ethics versus Law)

FACULTY

No. Correct	Facts	Missing Info	Ambiguous Info	Conclusion
0	5	19	20	15
1	5	9	9	14
2	19	1	0	0

STUDENT

No. Correct	Facts	Missing Info.	Ambiguous Info.	Conclusion
0	3	21	20	15
1	7	7	9	14
2	12	2	0	0

Assignment – Week 4

The next assignment was graded during week 4 of the semester. At this point, nearly 100% of the students are able to determine the facts. In addition, there is significant improvement in students ability to determine what information is missing that would be helpful in answering their questions as well as what information is ambiguous. Over 50% of the students are able to articulate at least one piece of missing information and one piece of ambiguous information that would help them reach a conclusion. However, this does not translate into students being able to put that information into their conclusion. At this point, students are not able to synthesize the information they have obtained so as to formulate a conclusion.

Topic: Procedure v. Substance (Judicial Process)

FACULTY

No. Correct	Facts	Missing Info.	Ambiguous Info.	Conclusion
0	0	4	7	11
1	1	17	16	12
2	22	2	0	0

STUDENT

No. Correct	Facts	Missing Info.	Ambiguous Info.	Conclusion
0	0	6	9	13
1	1	15	12	10
2	22	2	0	0

Assignment – Week 8

Week 8 was the next assignment that was graded. On this fact pattern, determining missing information proved to be a problem. However, 50% of the students were able to articulate material that was ambiguous. In reviewing the grading it appears that what was scored as missing information was often referred to by students as an ambiguity. What was scored as ambiguous information was referred to by students as information that was missing. Fewer than

50% of the students were able to reach a conclusion using either missing facts or ambiguous information.

Topic: First Amendment & Business (United States Constitution)

FACULTY

No. Correct	Facts	Missing Info.	Ambiguous Info.	Conclusion
0	2	15	9	20
1	6	4	13	2
2	16	5	2	2

STUDENT

No. Correct	Facts	Missing Info.	Ambiguous Info	Conclusion
0	0	14	6	15
1	7	5	12	7
2	16	5	6	2

Assignment – Week 12

The last assignment for the semester shows that nearly 90% of the students are able to identify at least 1 fact, with over 70% able to identify 2 facts. At this point 65% of the students are able to identify at least one piece of missing information that would help them in determining their answer. 50% of the students are able to identify one piece of information that is ambiguous, an answer to which would be helpful in answering the question. In terms of faculty grading, 50% of the students were able to use the information in reaching a conclusion. However, based on the TA's grading, fewer than 50% were able to reach a conclusion based on the missing and/or ambiguous information.

Topic: Guns (Product Liability)

FACULTY

No. Correct	Facts	Missing Info.	Ambiguous Info.	Conclusion
0	3	9	13	13
1	4	15	13	12
2	19	2	0	1

STUDENT

No. Correct	Facts	Missing Info.	Ambiguous Info.	Conclusion
0	4	9	11	16
1	6	14	14	8
2	16	2	0	2

CONCLUSIION

The study began with two hypotheses. The first hypothesis was that a methodology could be developed to teach specific critical thinking skills within the substantive course material in an

introductory business class. The second hypothesis was that this methodology could be assessed in an objective manner through the use of a grading rubric. The study indicates that there was some success in developing the methodology for teaching specific critical thinking skills. There was significant improvement in the ability of students to identify missing and ambiguous information that would be needed in reaching a reasoned conclusion although this ability did not translate into the reaching of a conclusion itself. It appears that the ability to use these inferences is a separate step – separate and apart from the ability to identify specific elements.

An unexpected positive that derived from the use of this method was an increase in students discussing differences of opinion in a forceful, professional manner. This is based solely on anecdotal information as the professor. As the semester progressed, class discussion became increasingly animated as different students put forth different opinions based on the information they needed and clarification on material they found ambiguous. Participation was not limited to a few students. The amount of discussion was such that it had to be given a specific time frame because of time constraints.

A fact pattern was developed for each chapter. However, not all of the material was able to be equally successful as a case problem. Topics such as torts lent themselves successfully to case problems. However, topics such as legal systems proved difficult to fit into the model. In the future, fewer case studies should be assigned – perhaps on half of the chapters. This would have a two-fold benefit. The cases themselves would be on topics suitable for this type of exercise. There would be less time spent on the case studies so that discussion would not have to be so time limited.

The second hypothesis was not as successful as the first hypothesis. The assessment mechanism appears flawed. Although two separate graders were used, using the same answer grid, differences appear in nearly each assignment. This is an indication of how difficult it is to properly assess critical thinking skills. Some thoughts on improving the assessment mechanism would be to have graders who did not participate in the actual course, either as a professor or as a teaching assistant. If the answers were made by one party and the grading done complexly separately, there might be a more consistent outcome. Another assessment option that could be used to determine whether or not a skill had been learned is to use a standardized test [13] given at the beginning and end of the semester.

REFERENCES

[1] Association to Advance Collegiate Schools of Business. Eligibility Procedures And Accreditation. 2011. (www.aacsb.edu/accreditation)
[2] Bloom, B. Taxonomy of Educational Objectives: the Classification of Educational Goals. New York, NY: Longman, Green and Company, 1965.
[3] Braun, N. *Critical Thinking in the Business Curriculum.* Journal of Education for Business, 2004, 79, 232-6.
[4] Camp, J. & Schneider, A. *Using Debate to Enhance Critical Thinking in the Accounting Classroom: The Sarbanes-Oxley Act and U.S. Tax Policy.* Issues in Accounting Education, 2010, 25, 655-75.

[5] Carrithers, D., Ling, T. & J.Bean. *Messy Problems and Lay Audiences: Teaching Critical Thinking Within the Finance Curriculum.* Business Communications Quarterly. 2008, 71, 152-70.
[6] Currivan, G. *Case Against Federal Aid to Colleges is Presented in Special Report.* The New York Times, May 18, 1959, E13.
[7]Kelley, C., Tong, P. & Choi, B. *A Review of Assessment of Student Learning Programs at AACSB Schools: A Dean's Perspective.* Journal of Education for Business. 2010, 85, 299-306.
[8] Peach, B., Mukherjee, A. & Hornyak, M. *Assessing Critical Thinking: A College's Journey and Lessons Learned.* Journal of Education for Business. 2007, 82, 313-20.
[9] Roever, C. *Using the Wall Street Journal to stimulate Critical Thinking.* Business communication Quarterly. 1998, 61, 66-70.
[10] Schott Karr, Susan. *Critical Thinking: A Critical Strategy for Financial Executives.* Financial Executive, 2009, 58-60.
[11] Sormunen, C. & Chalupa, M. *Critical Thinking Skills Research.* Journal of Education for Business, 1994, 69, 172-8.
[12] Wang, S. & Wang, H. *Teaching Design Thinking Through Case Analysis: Joint Analytic Process.* Decision Sciences Journal of Innovative Education, 2011, 9, 113-8.
[13] Watson, G. & Glaser, E. Watson-glaser Critical Thinking appraisal: Manual Forms A & B. U.S.A: The Psychological Corporation: A Harcourt Assessment Company, 1980.
[14] Winstead, K. *Marketing Debates: In the Classroom and Online.* Marketing Education Review. 2010, 20, 77-82.
[15] Zirulnick, A. *West Loses Edge to Asia in Education: Top Five OECD Findings.* The Christian Science Monitor. December 7, 2010, 1.

BLENDED AND ONSITE LEARNING ENVIRONMENTS: A COMPARISON OF DELIVERY MODELS AND COURSE OUTCOMES

Bhupinder S. Sran, Ph.D.
DeVry University
630 US Highway One
North Brunswick, NJ 08902
bsran@devry.edu
(732) 729-3962

John W. Weber, D.B.A
DeVry University
630 US Highway One
North Brunswick, NJ 08902
jweber@devry.edu
(732) 729-3900

ABSTRACT

With the emergence of online learning, the blended delivery model has also emerged. This modality combines onsite and online delivery. The goal of blended learning is to combine the best of each delivery method to provide the students with an enhanced learning experience. This paper will examine the difference in student outcomes between students enrolled in a blended version of a senior-level project management course to those who were enrolled in an onsite version of the course at a private college in New Jersey. A comparison for exam grade averages, final course averages and student course ratings will be presented.

BACKGROUND

Prior to the widespread adoption of the Internet, the most common form of education occurred in a face-to-face environment. In this environment, faculty and students would come together in a classroom to engage in the learning process [3]. However, since the early 2000's, enrollments in online courses and educational programs have significantly increased. The most common form of online instruction is considered to be asynchronous learning, where the students participate in computer mediated instruction to complete class assignments, engage in threaded discussions, and participate in other computer-based activities. One of the key benefits of this learning modality is that it provides learning opportunities to students who may not otherwise have access to education [1].

In recent years, education has not just been limited to onsite (face-to-face) and online deliveries. The advent of blended or hybrid courses has also emerged as a viable alternative for students. Blended courses combine both the onsite with online course delivery to provide students with a learning experience that capitalizes on the benefits, while also minimizing the limitations of each learning modality [1][2]. Colucci and Koppel (2010) indicated that onsite courses often provide for higher student motivation than online courses, as students better identify with their professor, fellow students, and course material. In contrast, students benefit from online coursework, as they have greater flexibility in pursuing their coursework, as they are less bound by time constraints [1]. The main time constraint relates to having to attend a class at a specific day and time. Blended courses have the potential to provide for an optimal balance of these benefits while providing more flexibility than onsite course by reducing the class meeting time or the number of classes.

In addition, Du (2011) posits that blended learning provides the opportunity to enhance student learning. One key benefit is that the online portion of a course can allow students to master the lower levels of learning prior to coming to class. These lower levels are considered to be the knowledge and comprehension levels of Bloom's Taxonomy. Then, the students could master the upper levels of learning in the face-to-face environment via faculty-guided activities [2].

Napeir, Dekhane, and Smith (2011) have highlighted that blended learning approaches can also benefit students that are lacking in computer literacy skills. The blended environment helps students overcome these limitations by providing opportunities for students to navigate the online environment through onsite orientations. In addition, blended learning provides opportunities for students to engage in activities that can be best experienced in the face-to-face environment [4].

The research provides insights that blended learning can deliver allows students to enhance their learning by engaging students in varied learning activities and modes. While it can provide the best of both worlds so to speak, it is most important to determine whether blended learning can provide for better student learning outcomes as compared to either the onsite or online modalities alone. There were several studies completed that examined student outcomes, which are presented here.

Colucci and Koppel's (2010) study examined student performance between one face-to-face and one blended section of a computer applications course. This course introduced the students to the Microsoft Office 2007 suite of business applications. Both groups had similar traits in terms of previous coursework, number of courses taken during the term, and number of work hours during the semester. The students' grades were compared for three homework assignments and two case studies, which were delivered online to the blended-learning group, and completed in class for the onsite group [1].

The results of Colucci and Koppel's (2010) study found that there was no statistically significant difference between the grades of the two groups. They also had examined other areas of student performance, such as missed assignments and completion of tutorials. Their results found that more of the students in the blended group missed assignments, and completed fewer tutorials, which did provide statistically significant results [1].In summary, this study found that there was not significantly different academic performance between the two groups.

Du (2011) also performed a study that compared student performance in an onsite, and a blended learning course. This particular course was Accounting Principles. This study examined student's performance in the students' overall, final exam, and homework. While the student performance for the overall grade was higher for the blended course (76.335) as compared to the onsite course (75.150), the results were not statistically significant. In addition, the results for the final exam, homework, and quizzes were all higher for the blended group versus the onsite group, but the difference in scores was also not statistically significant [2]. Napier, Dekhane and Smith (2011) have also found that researchers at the University of Central Florida found that the results of blended coursework have outcomes that are similar to those in face-to-face courses [4].

While student outcomes may not be significantly different between onsite and blended courses, there are other factors that need to be considered. One such factor relates to student perceptions and student satisfaction. Napier, Dekhane and Smith's (2011) study analyzed student responses

on end-of-term course evaluations. Their review of the evaluations found that the students preferred the flexibility of the online course, as they did not need to attend the onsite class as frequently. In addition, the students also liked having an onsite meeting, as it provided them with an opportunity to interact with their professor. Third, the students indicated that they assumed more responsibility for their learning, and also indicated that this responsibility better prepared them for being a productive member of the workforce [4].

However, not all of the feedback that Napier et al. (2011) received was positive. Several students indicated that blended learning required more of their time than a traditional course. In addition, those students with lacking time-management skills found that the blended methodology did not allow them to succeed [4]. Nonetheless, the student's feedback on blended learning was very positive.

While student feedback regarding blended learning has been positive, the studies presented above have found that there is no statistically significant difference between student performance in blended and onsite classes in the business and computer applications disciplines.

METHODOLOGY

This study assessed student performance in a project management course at a private college in New Jersey. The project management course is generally taken by seniors at the college. This course has been taught in various modes, including blended and onsite.

The onsite version of course met two times each week, for either 8 weeks or 15 weeks. The total meeting time was approximately 60 hours for the entire course. The final grade in the onsite version of the course consisted of two exams, a course project and daily work (i.e. homework and class exercises). The course material was delivered in face-to-face lectures. An eCollege shell was used to make announcements, record grades, and distribute documents, such as the syllabus.

The blended version of the course met once a week for 8 weeks, with each meeting lasting 4 hours. The total course meeting time was approximately 32 hours. An eCollege shell was an integral part of the course. In addition to using it for making regular announcements, recording grades, and distributing documents, it was also used for carrying out threaded discussions and compiling a webliography. Students were expected to post to the threaded discussions on at least three different days during the week.

This study compared student outcomes from three blended sections with three onsite sections of the project management course using exam grades, final course averages, and student evaluations. All sections were taught by the same instructor over a period of three years, from 2004 to 2007.

Students in all sections took two exams, which were similar in terms of content and difficulty. Their average on these two exams was used as the key variable for measuring student outcomes. Because of differences in the design of onsite and blended courses, the final course averages were based on different components. For onsite sections, the course average consisted of grades for exams, daily work (e.g. homework and class exercises), and a group project. Final averages

for blended sections included all of these elements and, in addition, online work (e.g. threaded discussions and webliography).

Data about grades was retrieved from the online grade book for each section of the course. Students who did not attend class and did not do any of the assignments were removed from the data. To minimize the effect of outliers, the students with the highest and the lowest exam averages in each group were also removed from the data set. After this process, there were a total of 57 students in the blended group and 59 students in the onsite group.

The course evaluation data for the sections under study was also retrieved. Evaluations for two of the onsite sections were not available as these were done on paper and records were not retrievable. For the other section, only the class rating summaries were available. Individual student evaluations were not available.

Statistical methods were used to perform the analysis for this study. These included use of basic descriptive statistics and two-sample t-tests. Minitab was used for carrying out all statistical analysis.

FINDINGS AND DISCUSSION

In this section, the results of the statistical analysis are presented. As mentioned earlier, exam averages, final averages, and student feedback are examined.

Exam Averages

The table below summarizes the descriptive statistics for exam averages.

Exam Averages – Descriptive Statistics

Delivery Mode	N	Mean	StDev	Minimum	Median	Maximum
Blended	57	77.77	8.36	53.50	79.50	94.00
Onsite	59	78.84	10.38	57.50	80.50	95.50
All	116	78.32	9.41	53.50	79.75	95.50

The mean of the exam averages for students was 77.77 in the blended sections and 78.84 in the onsite sections. A two-sample t-test was performed to determine whether the difference between the two means was significant. This test yielded a t-score of -0.61 with a p-value of 0.543, thus indicating that the difference in the means of the two groups was not significant.

Final Course Averages

The final course averages for both blended sections and onsite sections are presented in the table below:

Final Course Averages – Descriptive Statistics

Delivery Mode	N	Mean	StDev	Minimum	Median	Maximum
Blended	57	81.88	9.25	53.11	83.11	95.78
Onsite	59	86.00	6.40	62.50	86.84	96.75
All	116	83.97	8.16	53.11	84.71	96.75

The final course averages seemed to vary more than the exam averages. The mean final course average for the students in blended sections was 81.88, whereas the mean for onsite sections was 86.00. A two-sample t-test produced a t-score of -2.79 and p-value of 0.0006, indicating that there was a significant difference in final course averages. These results should be treated with caution since the components and the weights used to determine final grades varied in the two delivery modes. For example, in onsite sections, greater weight was assigned to daily work (e.g. homework and class exercises), an area in which students do significantly better. In onsite sections, there were 9 to 10 daily assignments, whereas blended sections only had 4 such assignments. Also, the grades on the group course project were significantly higher for the onsite sections (mean=88.97) than for the blended sections (mean=84.05). There are several potential factors that may have caused this difference, one being the increased amount of time students had in class in the onsite sections to work on the project.

Student Feedback

Student course ratings for blended sections and onsite sections were nearly identical. Student course ratings were available for one of the onsite sections and all three blended sections. Twenty-two students in the onsite section completed the course evaluation with a mean course rating of 3.68 (out of 4). Although data was not available for all onsite sections, the rating in this section is very close to the ratings that this instructor has received in the past. Fifty-one students in blended sections completed a course evaluation with a mean course rating of 3.66. Qualitative student feedback was not available.

CONCLUSION AND LIMITATIONS

Student performance on exams is consistent with previous research findings which indicate that there is not a significant difference in student outcomes in onsite and blended environments. A comparison of student course ratings also is in agreement with past research. However, this study did find a significant difference in final course averages for the two delivery modes. The students in onsite sections had higher final course averages than students in blended sections. Initial exploration of the data indicates that these higher course averages may be the result of higher grades on group course projects in onsite sections. It appears that the group projects benefitted from increased face-to-face student interaction provided by onsite courses. This finding suggests that faculty should look for ways to further increase student interaction in a blended environment through strategic use of online tools such as group threaded discussions and web conferencing.

The results of the study should be interpreted with several limitations in mind. It is important to note that there were differences in the way the final course averages were calculated for students in the two groups. Since the study covered courses taught over a three-year time span from 2004 to 2007, there may have been changes in course content, pedagogy, and grading criteria during that time span. Also, any new strategies implemented in blended courses since that time were not considered in this study. Since this study focused on just one upper level business course taught by one faculty member, its findings should not be generalized to courses at other levels and in other disciplines.

REFERENCES

[1] Colucci, W. & Koppel, N. *Impact of the Placement and Quality of Face-To-Face Meetings in a Hybrid Distance learning course.* American Journal of Business Education, 2010, 3(2), 119-130.

[2] Du, C. *A Comparison of Traditional and Blended Learning in Introductory Principles of Accounting Course.* American Journal of Education, 2011, 4(9), 1-10.

[3] Gautsch, S. & Griffy-Brown, C. *The Perfect Blend. Biz Ed,* 2010, 9(6), 28-32.

[4] Napier, N. P., Dekhane, S. & Smith, S. *Transitioning to Blended Learning: Understanding Student and Faculty Perceptions.* Journal of Asynchronous Learning Networks, 2011, 15(1), 20-32.

COMMUNITY BASED SERVICE LEARNING

Kellyann Berube Kowalski, Department of Management & Marketing, Charlton College of Business, University of Massachusetts Dartmouth, 285 Old Westport Road, North Dartmouth, MA 02747, (508) 999-8327, kkowalski@umassd.edu

Jennifer Ann Swanson, Department of Business Administration, Stonehill College, 320 Washington Street, Easton, MA 02357, (508) 565-1349, jswanson@stonehill.edu

Jim Lee, Department of Business Administration, Stonehill College, 320 Washington Street, Easton, MA 02357, (508) 565-1478, jlee@stonehill.edu

Nadia Abgrab, Department of Business Studies and Economics, Salve Regina University, 100 Ochre Point Avenue, Newport, RI 02840, (401) 341-3209, nadia.abgrab@salve.edu

WORKSHOP FORMAT

The focus of this workshop is community based service learning, which can provide students with a very rewarding and rich learning experience. Four faculty from three different institutions of higher education will begin the workshop by briefly providing an overview of how they incorporate community based service learning into their marketing, management, and international business courses. Examples of these service learning activities can be found in Appendices A, B, and C. Appendix A describes service learning in a Building Leaders Learning Community at a private liberal arts college, Appendix B provides an overview of a Teams course that incorporates service learning at a state university, and Appendix C discusses how service learning is incorporated into an International Issues in Consumer Behavior course at a private liberal arts university.

The workshop will also address the following questions related to community based service learning:

- Should community based service learning teams be assigned by instructor, chosen by the students, or randomly assigned?
- What types of deliverables should be required for the service learning project?
- How can we ensure that students are "learning" and that course content is tied into the service project? For example, it has been found that service learning works best when there is a strong reflective component to the exercise – requiring that students make a cognitive link between service and learning. How can this be best accomplished?
- How should students be assessed? As individuals or teams? Can members of the same team get different grades? Should team members evaluate each other, themselves? Should clients evaluate teams? Do teams develop their own evaluation instruments?
- How can the connections with the community be made? How can student teams find clients in the community?
- Should teams individually meet with professor during the semester?

- Would a course website, such as Web CT, somehow be utilized to enhance the teams' experiences during the service learning project?

APPENDIX A
BUILDING LEADERS LEARNING COMMUNITY

The Building Leaders Learning Community links the concepts of adolescent developmental psychology with organizational behavior and uses the problem context of mentoring at-risk teenagers to enable student teams to apply the content of both courses to their client, as they work to provide direct services for youth of their own age, and to their team development as they strive to coordinate their efforts in teams of 13 students. Concepts in adolescent development include cognition and learning, moral development, and issues relating to risky behavior during the adolescent years which relate both to the students themselves and to their clients. Concepts in organizational behavior relate to the building of effective cooperative teams, leadership, conflict resolution, personal responsibility, and ethics in organizational contexts. Most client organizations deal with at-risk youth who have dropped out of high school and who are attempting to prepare to take the GED examination, or who are at risk to drop out and are being supported in their efforts to remain in school. The campus has a center for community based learning whose staff members obtain permission from area service providers, whose mission would be facilitated by student team projects, to allow a team to assist with a particular problem faced by the client organization. Teams are responsible for negotiating the scope of the project, project design, implementation, outcome assessment, and for formally reporting their results. Pedagogical structure involves weekly meetings of teams to share results, to make a series of reports that are designed to facilitate the development of internal team structure, and to provide a block of time to enable student teams to interact within their academic schedules (one 150 minute meeting per week). Deliverables include a team report on outcomes to be delivered to the client, a formal 20 minute executive summary of outcomes provided to a joint meeting of client representatives and all students, and an individual reflection paper where students use the theories of both classes to assess on their academic achievement and personal development as a result of the course.

APPENDIX B
DEVELOPING AND MANAGING WORK TEAMS

As the instructor for the course, "Developing and Managing Work Teams," I have focused the course around a significant community based service learning project. The course, which is an exploration of the concept of team development in U.S. organizations today, is a natural fit for a service learning component. Students in the course examine the approaches, successes, and failures of team development, and are be required to engage in team management and development efforts. The team based service learning project is a pedagogy that enables students to meet those course objectives.

Depending on the semester the course is taken, students work on one of two community based service learning projects. During the fall semester, student teams work on an "America Recycles Day" project. All teams work with our District Recycling Coordinator to increase recycling in the local community. Each team is responsible for finding and working with a second community partner to coordinate an event or events that increases recycling awareness, education, and behavior in the local community. During the spring semester, student teams work on a "People Powered Float" project. All teams must participate in a local Earth Eve parade by building and entering a float. Each team is responsible for finding a community partner that they will serve in some capacity (e.g. volunteering), as well as represent and promote with their float. Ultimately my goal is to uncover the necessary variables to best utilize service learning and community service in developing our future leaders and effective team members. Learning to be an effective member of a team is an action-oriented process. Toward that end, this service learning project provides an outlet where teams develop and carry out a project that makes a significant difference in the community. The service learning project provides students with an opportunity to practice and reflect on their teamwork concepts and skills.

APPENDIX C
INTERNATIONAL ISSUES IN CONSUMER BEHAVIOR

In the International Issues in Consumer Behavior course, students have the option of earning an additional one service learning credit by working on a community need-related non-profit project. This option is only awarded to the students who successfully link theoretical values to practical community outreach projects. Students are asked to work together as marketers to create interactive demonstrations, present their findings, and teach the course concepts to the external target market in an exciting and fun manner. At the conclusion of their service experience, students are asked to write up a paper that reflects on the purpose, usefulness, and meaning of their contributions to society.

One example of this community service learning experience involved students working on interactive projects with 6^{th}, 7^{th}, and 8^{th} grade students at a local school in Rhode Island. These projects focused on eight countries (Brazil, Cambodia, China, Greece, Ireland, Italy, Japan, and Portugal) with the goal of teaching others how market economics and culture findings impact the lives of international consumers. Various marketing strategies were implemented to teach about the range of meanings consumers attach to the subject of 'health and wellness' on a global scale. Primary research enhanced the secondary searches to define behavior of the target market/country, competing brands, packaging, usage rates, and physical attributes of product/service: color, size, smell, and texture.

Another example of a community based service learning project in the course was an exploratory research activity focused on four fastest growing consumer markets in the world: Brazil, Russia, India, and China (BRIC). This project evaluated key consumer characteristics that world-wide businesses must analyze as they expand into emerging markets. Traditions, trends, and needs of individuals living in countries with cultural challenges were some of the factors explored. Using the consumer behavior topics presented in class, an interactive presentation was made to a 12-14 year old target market at an area school – to teach them about the values, rituals, heroes, and symbols of these countries. Through a hands-on career-fair setting, a fun educational platform was created to demonstrate how multinationals could generate revenue for each country-- thus allowing for a better quality of life.

THE PERCEIVED PERFORMANCE OF WESLEY COLLEGE STUDENTS IN A BASIC MATH COURSE

Derald E. Wentzien, Ph.D, Wesley College
Carle Ax, Wesley College

ABSTRACT

The perceived performance of Wesley College students enrolled in a developmental math course was measured and compared to their actual performance. A survey was administered to the students to identify their perceived current grade at the time of the survey. The instructors in the classes were also asked to identify the actual grade of each student. A Paired-Sample Sign Test was performed to determine if the students were correctly identifying their actual performance. The results indicated that the students did not correctly perceive their true performance; instead the students overestimated their grades. In particular, the differences between perceived and actual grades increased as the actual scores decreased.

BACKGROUND

Wesley College is a small liberal arts college located in Dover, DE with an enrollment of approximately 1,500 undergraduate students. Wesley College is accredited by Middle States, and retention and graduation rates of the Wesley College students are reviewed during the 10-year re-evaluation process. Over the last few years approximately 25% of incoming freshman students took a basic math skills course and a significant percentage of them did not pass the class. This poor performance in the developmental math course significantly lowers the retention and graduation rates of students at Wesley College.

Assessment and scoring rubrics have taken on an important role in recent years. Many people have stressed the need for accurate and timely assessments [1] [2]. The faculty at Wesley College has adopted assessment tools in the last few years in the hope that retention and graduation rates will improve. The results of the assessment have been enlightening, but there still seems to be misunderstanding by the students as to their performance. Many students have indicated that they do not understand why they failed a particular class. This study was initiated to begin to study some of the causes of these poor performances. In particular, the faculty in the math department wanted to determine if the students accurately perceived their true performance in the developmental math class. In other words, do the students who fail the course know they are failing the class before they receive their final grades?

METHODOLOGY

This study was conducted to determine if the students enrolled in MA099, Basic Math Skills, knew how well they were actually doing in their math class. Eleven sections of MA099 (Basic Math Skills) were surveyed during the month of November 2010. By collecting the data during the month of November, it gave students enough time (roughly two months) to hand in enough assignments and tests to get a good idea of their current grade. A total of 117 surveys were collected out of a total of 176 enrolled students in the MA099, Basic Math Skills courses. This represented a total response rate of 66.5%. The students were presented with a survey before their class started and then their teachers provided each student's numeric average to the researcher. All the students' names were taken out and changed to Student 1, Student 2…etc. so as to respect and protect their privacy.

Estimates were created and used to facilitate comparisons between perceived and actual letter grades and numeric averages. The estimate of the students' perceived numeric scores and the determination of the students' actual letter grades needed to be determined before the comparisons could be made. The calculations of the letter grades were straight forward; the grading criteria from the syllabi were used to convert the numeric scores to letter grades. The students perceived letter grade was converted to a numeric score so that hypothesis testing could be performed on the significance of any differences. Asking the students to provide a perceived numeric average was not so as to simplify the process for the students. It was also believed that a more realistic perception could be obtained by asking for the letter grades instead of averages. The perceived letter grade was compared to the numeric score provided by the instructor to determine a perceived numeric average. If the student perceived a C and their average was 75.6%, a perceived grade of 75.6% was used to show that the student accurately perceived their letter grade. If the student perceived a letter grade of B and their actual average was 75.6%, a perceived grade of 80.0% was used. 80.0% was used since it was the closest score in the B range (80.0% to 89.9%) to the student's actual average. If the student perceived a letter grade of D and their actual average was 75.6%, a perceived grade of 70.0% was recorded. The 70.0% was used since it was the closest score in the D range (60.0% to 69.9%) to their actual average.

After the data was collected and the student names hidden to protect their identity, a variety of statistical analyses were then conducted to study the results. A pivot table was created in Excel to compare the actual letter grade with the perceived letter grade that the student believed he or she had at the time of the survey. The pivot table was constructed to identify whether a linear relationship, a lower triangle relationship, or an upper triangle relationship was readily discernable. In an ideal world, the information would have a linear relationship; that is the students' perceived letter grade was exactly the same as their actual letter grade. A lower or upper triangular relationship would identify a tendency for students to overestimate or underestimate their performance.

A goodness of fit test was used to determine if the perceived and actual numeric data was normally distributed over the range of data. If the perceived and actual grades were normally distributed a paired t-test would be used. The goodness of fit test indicated that the perceived numeric scores were not normally distributed, therefore, a non-parametric paired sample sign test was used to determine if there was a significant difference between the perceived and actual

numeric averages. A total of five paired sample sign tests were performed, one for the entire sample and separate tests for each perceived grade level (A, B, C, D, and F).

RESULTS

The results indicated that there is a misperception between the perceived letter grades and the actual letter grades. Table 1 shows a pivot table between the perceived letter grades and the true letter grades. All of the values would lie on the main diagonal of the pivot table if students correctly perceived their actual grade. Instead, the data is lower triangular, meaning that several students over-estimated their actual performance. In some cases the over-estimation was significant; of the failing students, three believed that they were earning a B, six believed they were earning a C, and five believed they were earning a D. Only six of the twenty failing students believed that they were failing. A total of fifteen out of twenty-four students earning a D thought that they were doing better; one thought that they were earning an A and three thought that they were earning a B. Thirteen out of thirty-three students earning a C thought that they were doing better. In this case all but one thought that they were doing better by only one letter grade.

Table 1

	Perceived Letter Grade					
Actual Letter Grade	A	B	C	D	F	Total
A	13	2				15
B	7	17	1			25
C	1	12	20			33
D	1	3	11	9		24
F		3	6	5	6	20
Total	22	37	38	14	6	117

Hypothesis tests were then performed to determine if the differences between the perceived and actual numeric scores were significant. Table 2 shows a comparison between the perceived numeric averages and the actual averages, and whether the paired sample sign test yielded a significant difference at a 5% alpha level.

A significant difference was observed in the mean paired differences for the entire sample; this was the preconceived notion of the math faculty. On many occasions math faculty report that it is as though the students do not have a grasp of the gravity of their academic condition. The results of the hypothesis test confirmed the belief.

The students were segmented by their current grade and each group analyzed individually. Significant differences were observed in the mean paired differences for the students earning a C, D, or F. Smaller p-values were observed as the actual grades decreased. The results indicated that the students who are failing are more likely to misunderstand their current situation.

Table 2

Actual Letter Grade	Perceived Average	Actual Average	Significant Difference
Entire Sample	76.97	73.53	Yes
A	93.35	93.62	No
B	85.49	84.67	No
C	77.06	75.20	Yes
D	69.95	65.42	Yes
F	62.32	51.49	Yes

CONCLUSION

Since the results indicated that the weaker performing students (the students that Wesley College is not retaining) are not fully grasping their academic standing in MA099 Basic Math Skills, a follow up survey will be performed to determine exactly why the students are not aware of their current situation. The follow up survey will ask the students to identify what factors contribute to why they are currently earning the grades that they think they are earning.

REFERENCES

[1] Brookhart, S.M. (1999). *The Art and Science of Classroom Assessment: The Missing Part of Pedagogy*. ASHE-ERIC Higher Education Report (Vol 27, No. 1) Washington, DC: The George Washington University, Graduate School of Education and Human Development.

[2] Westat (2001). *ABC on Scoring Rubrics Development for Large Scale Performance Assessment in Mathematics and Science*. Posted online at http://lsc-net.terc.edu.

EXAMINING THE USE OF HOFSTEDE'S UNCERTAINTY AVOIDANCE CONSTRUCT IN ETHICS RESEARCH: A 29-YEAR REVIEW

Julia D. Davis
Gabelli School of Business
Roger Williams University
1 Old Ferry Rd., Bristol, RI 02809
jdavis413@g.rwu.edu
(401)254-3672

Richard A. Bernardi
Gabelli School of Business
Roger Williams University
1 Old Ferry Rd., Bristol, RI 02809
rbernardi@rwu.edu
(401)254-3672

Susan M. Bosco
Gabelli School of Business
Roger Williams University
1 Old Ferry Rd., Bristol, RI 02809
sbosco@rwu.edu
(401)254-3175

ABSTRACT

Our research will review articles in the Journal of Business Ethics that used Hofstede's cultural dimensions of uncertainty avoidance. Journal of Business Ethics is the foremost outlet for ethics research. Our goal is to identify research that used Hofstede's uncertainty avoidance construct in their literature reviews and/or their analysis. Consequently, our research indicates that approximately 84% of the articles used uncertainty avoidance strictly in the literature review and 16% used uncertainty avoidance as a research variable or to compute a variable. This review provides a foundation for future research using this construct.

KEYWORDS: Hofstede, Uncertainty Avoidance, and Ethics.

INTRODUCTION

This paper examines the use of Hofstede's cultural construct of uncertainty avoidance in international ethics research (although the majority of the research we examined in the course of this study included all of Hofstede's constructs). Hofstede [55, p.112] defined culture as a "*set of likely reactions of citizens with a common mental programming. . . . reactions need not be found within the same persons, but only statistically more often in the same society*". The uses of Hofstede's cultural dimension of uncertainty avoidance vary in the *Journal of Business Ethics* as well as the *International Journal of Value Based Management* and Teaching Business Ethics, which were merged into the *Journal of Business Ethics* in 2004. This paper breaks down the different uses of uncertainty avoidance into five distinct categories. The grouping of the different uses allows readers to quickly search and understand what information is available in these journals. A similar data compilation was performed for the *Journal of International Business* by Rapp et al. [106]. The five categories that we used to divide uncertainty avoidance are: definition use, brief comparisons, literature reviews, alternative cultural indices and as a variable. There were 19 articles found in these journals that used uncertainty avoidance as a variable. Of the 19 articles, 10 were human participant studies, six were country based studies and three articles fit into another category. It was found that uncertainty avoidance had many strong associations with other behaviors.

Articles that used uncertainty avoidance as a variable found that uncertainty avoidance has a positive association with worker rights, corporate social responsibility, good business, social desirability response bias, relativism, Organization for Economic Co-operation and Development *(OECD)* convention ratification, Corruption Perceptions Index, nationalism, power distance, ethical policies and ethical values. There are negative associations between uncertainty avoidance and individualism and practices of ethics. Six articles used uncertainty avoidance as a variable but to determine correlations between uncertainty avoidance and Chinese and US Caucasians, sexual harassment, whistle blowing, sensitivity and software piracy, but did not have significant results. The multiple uses of uncertainty avoidance helped make definitive associations or shed light onto an area that needs more research.

METHODOLOGY

Identifying Articles

The first step in the data gathering process was to identify articles to examine in the *Journal of Business Ethics,* which has been published since 1982. We began by using the journal's online search function for 'uncertainty avoidance'. To ensure that we had identified all articles, we also used *Google's Advanced Scholar* searching for 'uncertainty avoidance' in the *Journal of Business Ethics*. We included only original journal articles; book reviews, comments, discussions and rejoinders were not included in our analysis. While this journal has been published since 1982, the first use of Hofstede's uncertainty avoidance construct was in 1990; so our study actually includes articles between 1990 and 2010 (e.g., a 21-year period).

While our primary aim was to include articles from the *Journal of Business Ethics*, the publisher of this journal also published the *International Journal of Value Based Management* from 1988 through 2003 and *Teaching Business Ethics* from 1997 through 2003. These journals were 'merged into the *Journal of Business Ethics*' at the beginning of 2004. We submit that article counts in the *Journal of Business Ethics* between 1988 to 2003 (i.e., the time period when these journals were independently published) would be understated compared to other years if these journals were not considered in the article count. Consequently, our total article count and classifications include publications from the *International Journal of Value Based Management* and *Teaching Business Ethics*.

Of the 154 articles that included Hofstede's uncertainty avoidance, 133 were in the *Journal of Business Ethics;* 12 were in the *International Journal of Value Based Management*; and, nine were in *Teaching Business Ethics*. Of the 133 articles in the *Journal of Business Ethics,* the journal's search engine initially identified 130 articles using "uncertainty avoidance" as our search term. However, two of these articles were not used; one of these used the term uncertainty avoidance outside of the context of Hofstede and the other was a two-page issue introduction article that noted one of the featured articles had a significant finding using uncertainty avoidance as a variable. Consequently, our initial journal search yielded 128 articles; the remaining five articles in the *Journal of Business Ethics* were identified using the Advanced Google Scholar search engine; for some unknown reason, these articles are not in the *Journal of Business Ethics* online data base.

Classifying Articles

After identifying the articles that included Hofstede's uncertainty avoidance, the second step in the data gathering process was to determine how uncertainty avoidance was used in each article. We classified the 154 articles into five categories that include: as a definition, brief comparisons, literature reviews, alternative cultural indices and uncertainty avoidance as a variable. After classifying the articles, we subsequently reviewed the classifications for validation purposes and resolved any differences in classification.

RESEARCH FINDINGS

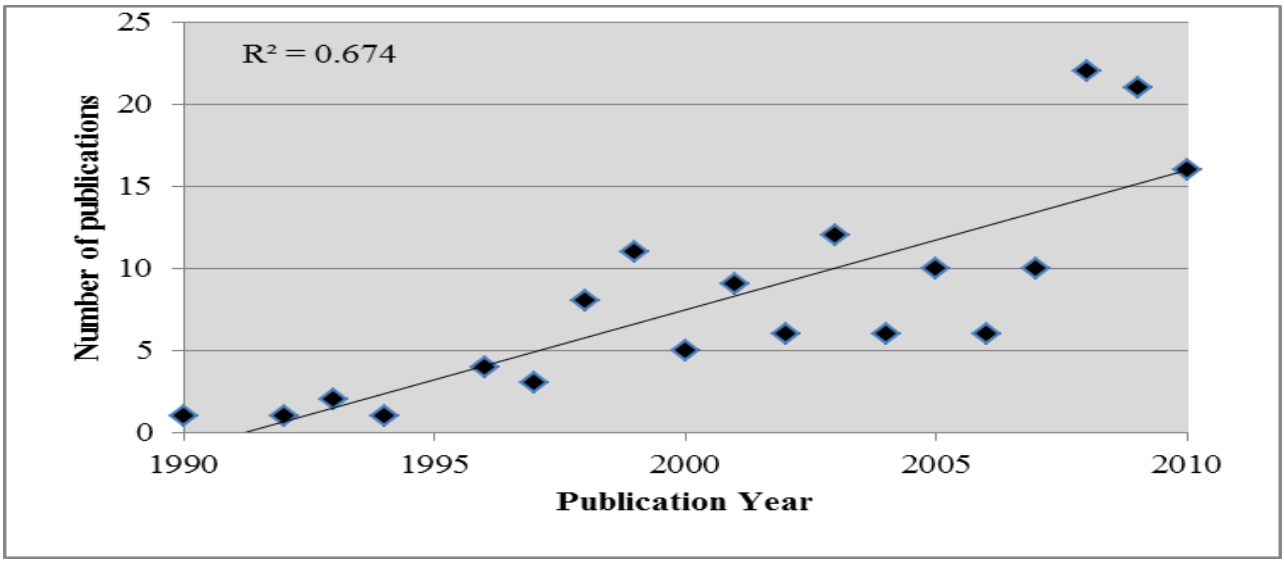
Overview

For the 25 years of this study, we found 154 articles. Figure 1 shows the increasing use of Hofstede's cultural dimension of uncertainty avoidance in research shown in the *Journal of Business Ethics*. Even though Hofstede's book first appeared in 1980 that cited four cultural dimensions including uncertainty avoidance, the first article found in the *Journal of Business Ethics* that referred to Hofstede's cultural dimension of uncertainty avoidance did not appear until 1990. There are multiple reasons for the time lag between Hofstede's first article and the first article referring to his cultural dimensions in this journal. One reason is that research at that time was not as readily accessible as it is today. In addition, there is the concern that articles before 1992 were scanned into online databases which disallow users to search articles for keywords. Therefore there could potentially be articles before 1992 that referred to Hofstede's cultural dimensions, but they did not appear in these searches. Since 1992, however the average use of uncertainty avoidance in the *Journal of Business Ethics* has increased by one article every

year. A univariate regression model used the number of years since Hofstede's cultural dimension of uncertainty avoidance was first cited in the *Journal of Business Ethics* and the number of articles every year to calculate an adjusted r^2 of .674 (Figure 1).

FIGURE 1

Number of Articles Using Hofstede's Uncertainty Avoidance



Due to the varied uses of uncertainty avoidance, the 154 articles that used uncertainty avoidance were categorized into five groups; each with subcategories. The first group includes 42 articles that include only definitional use of uncertainty avoidance as background information. In the second group, 39 articles use uncertainty avoidance as a brief comparison. Forty-eight articles are in the group section which uses uncertainty avoidance in the literature review. The fourth group includes four articles that use uncertainty avoidance in addition to an alternative cultural index. The fifth group includes 21 articles that use uncertainty avoidance as a variable in the research. The division into five groups allows readers to better organize the different uses of uncertainty avoidance. The number of articles appearing in our tables exceed the number of total articles in the study because four of the articles [150] [98] [113] [75] appear in more than one table.

Definition Articles

Table 1 organizes 42 articles that utilize a definition of uncertainty avoidance. The table includes articles from 1990 to 2010. The number of times uncertainty avoidance was cited briefly remained constant except for 2008 and 2010. The four panels in Table 1 show the subcategories of uses of uncertainty avoidance. Four articles in Panel A define uncertainty avoidance. The definitions of uncertainty avoidance were used to explain corruption [13] [34] management [95] and ethics [107].

TABLE 1

Mention of Uncertainty Avoidance

Panel A: Defines Uncertainty Avoidance

Author(s)	Year	Author(s)	Year
Beets	2005	*Parnell & Hatem*	*1997*
Collins et al.	2009	Rawwa et al.	1998

Panel B: Defines Uncertainty Avoidance along with other cultural dimensions

Author(s)	Year	Author(s)	Year
Chandler & Graham	2010	Lin	2008
Clikeman et al.	*2001*	Primeaux et al.	2003
Cohen & Single	2001	*Robertson & Schlegelmilch*	*1993*
Dunn & Shome	2009	*Sama & Papamarcos*	*2000*
Hay et al.	*2001*		

Panel C: Defines Uncertainty Avoidance and lists other cultural dimensions

Author(s)	Year
Singhapakdi et al.	2001
Tang	2008

Panel D: Lists Uncertainty Avoidance along with other cultural dimensions

Author(s)	Year	Author(s)	Year
Davis et al.	1998	MacNab & Worthley	2008
Aygün et al.	2008	*Nebenzahl et al.*	*2001*
Cheung & Chan	2005	Ng & Burke	2010
Chiu et al.	2008	Peterson et al.	2010
Erben & Güneşer	2008	Phau & Kea	2007
Fam & Waller	2003	Resick et al.	2006
Jin & Drozdenko	2010	Robertson & Fadil	1999
Jin et al.	2007	Schwartz	2009
Jing & Graham	2008	Sims	2002
Jones	1999	Singhapakdi & Vitell	1999
Kim	2003	Sison	2000
Lämsä et al.	*2000*	*Sotto & Kohls*	*1990*
Lin	2010	Zgheib	2005

UA Uncertainty Avoidance
Italicized data indicate publications in the *International Journal of Value Based Management* and *Teaching Business Ethics*.

In the following subcategory, Panel B, uncertainty avoidance is defined along with other cultural dimensions. The information provided about uncertainty avoidance, and more so, Hofstede's cultural dimensions, provided a basis for understanding differences in cultures. These differences

in cultural values were used to explain varied beliefs related to advertising [75], corruption [22] earnings management [31] ethics [14] [53] [112] flexible work arrangements [32] and justice [103] [114].

Of the articles in Table 1 in Panel C, two articles also list Hofstede's cultural dimensions along with uncertainty avoidance. Tang [139] explains the how uncertainty avoidance affects people's contribution to the public well-being and Singhapakdi et al. [127] uses uncertainty avoidance to explain its potential effects on marketing ethics. Both articles define and list Hofstede's cultural dimensions to explain individuals' behavior within certain cultures.

Panel D within Table 1 contains 26 articles that list uncertainty avoidance with the other Hofstede cultural dimensions. Thirteen articles included in Panel D listed uncertainty avoidance with Hofstede's other cultural dimensions; the uses of uncertainty avoidance in these articles varied. The thirteen articles used the list of cultural dimensions to explain how culture can affect behaviors. In several areas, four articles [38] [89] [100] [158] observed the association between uncertainty avoidance and ethics. Nine articles found relationships between uncertainty avoidance and advertising [44], business education [71] corruption [118] deception within the workplace [125], leadership [26], paternalism [42], peer group relations [132], regulation policy [62] and whistle-blowing [83].

There were other uses of the construct within Panel D. Three articles [8] [90] [110] had uncertainty avoidance listed because another cultural dimension (not uncertainty avoidance) played an integral role in the research and the authors found it necessary to mention the other cultural dimensions. Four articles [28] [61] [79] [101] did not use Hofstede's cultural dimensions in the research, but included the cultural dimensions within the limitations section of the paper. Another four articles [63] [69] [126] [129] related the ability to understand countries' values to observing cultural dimensions and listed the available dimensions by which one could judge a county

Uncertainty avoidance was used differently in the final two articles. One used uncertainty avoidance to cluster similar countries [109] and the other to reflect on previous research [60]. The listing of uncertainty avoidance with the other Hofstede cultural dimensions allowed researchers to explain associations, reflect on research limitations, understand how to evaluate countries and illustrate uses of other cultural dimensions.

Brief Comparison Articles

The articles in this group only provided a brief mention of uncertainty avoidance in an article. These comparisons using uncertainty avoidance are separated into three subcategories in Table 2: single or multiple country studies (Panels A and B), action or belief articles (Panel C), and miscellaneous uses (Panel D). A breakdown of the number of articles in each of these three categories includes: uses of uncertainty avoidance to compare countries (22 articles), uncertainty avoidance association to another behavior (14 articles), and miscellaneous uses (three articles).

TABLE 2

Using Uncertainty Avoidance in brief comparisons

Panel A: One-country studies

Author(s)	Year	Theoretical Premise
Sidani & Thornberry	2010	Arab Societies - Mentions previous research that stated had high UA ratings
Al-Khatib et al.	*2004*	*Arab-Gulf Culture – Mentions UA rating*
Lowe	2008	China - Mentions UA rating
Tan	2008	China - Mentions UA rating
Sidani & Jamali	2010	Egypt – Mentions UA rating
Monga	2007	India - Uses UA to explain behaviors of Indian Managers
Brinkmann	2002	Norway – Listed UA Score with other Hofstede cultural dimensions ranks
Atakan et al.	2008	Turkey – Mention UA rating
Leung et al.	2009	Thailand - Mentions UA rating
Jose & Thibodeaux	1999	US - Used UA to explain behaviors of American Managers
Simga-Mugan et al.	2005	US - Uses UA to explain behaviors of Americans

Panel B: Two-country studies

Author(s)	Year	Theoretical Premise
Wimalasiri	2004	Fiji and India - used UA to explain why they avoid risk and have measures doing so
Chhokar et al.	*2001*	*France and Russia –used UA to compare*
Whitcomb et al.	1998	Hong Kong and Taiwan - defines UA and uses UA to compare
Wan et al.	2009	Hong Kong and Taiwan - suggests UA was the only major cultural difference
Marta et al.	*2003*	*US and Egypt – uses UA to compare*
Burton et al.	*2000*	*US and Hong Kong – used UA to compare*
Marta et al.	2008	US and Mexico - mentions the UA differences
Smith & Hume	2005	US and the Netherlands - uses UA to briefly state a difference
Brody et al.	*1999*	*US and Taiwan –uses UA to compare whistle-blowing differences*
Singhapakdi, et al.	2008	US and Thailand - used UA to compare
Leong et al.	2004	US and UK - used UA to compare

Panel C: Uses Uncertainty Avoidance as it associates with an action or belief

Author(s)	Year	Suggests a positive, negative or possible association with UA
Cherry	2006	A positive association between UA and lower ethical problems
Haswell et al.	*1999*	*A positive association between UA and more rule specific control systems*
Li et al.	2009	A positive association between UA and technology uncertainty

O'Higgins & Kelleher	2005	A positive association between UA and ethics
Preble & Hoffman	1999	A positive association between UA and a rules orientation
Wang & Lin	2009	A positive association between UA and high brand loyalty
Vitell et al.	*1993*	*A positive association between UA and intolerance of deviations from group norms*
Moores	2008	A negative association between UA and software piracy - listed all cultural dimensions
Patel & Schaefer	2009	A negative association between UA and ethical decision making
Sood	*1998*	*A negative association between UA and religious factors*
Zhang et al.	2009	A negative association between UA and whistle-blowing in China
Fleming et al.	2010	A possible association between UA and moral reasoning - listed all cultural dimensions
Mills et al.	2009	A possible association between UA and Conformity.
Sanchez et al.	2008	A possible association between UA and corruption

Panel D: Miscellaneous uses of Uncertainty Avoidance

Author(s)	Year	Theoretical Premise
Bowen	2007	Mentions UA use in comparing firms
Insch et al.	2008	Uses UA & Masculinity to relate to female expatriates
Khatri & Tsang	2003	Suggests using UA in future research

For 22 articles, uncertainty avoidance was applied at a country level to explain actions or beliefs. The countries included in these articles vary. Panel A contains articles that focused on a single country research. These included Arab societies [121], Arab-Gulf culture [2], China [81][138], Egypt [120], India [87], Norway [17], Turkey [5],Thailand [73] and the United States [65] [122]. In Panel B, the articles that reviewed two countries compared Fiji to India [155], France and Russia [27], and Hong Kong to Taiwan [151] [154]. In addition to those countries, the United States was compared to Egypt [84], Hong Kong [20], Mexico [85], the Netherlands [130], Taiwan [18], Thailand [128] and United Kingdom [72]. The articles in this section used uncertainty avoidance to show cultural differences to provide background information or reasoning for future discussions.

Associating uncertainty avoidance and behaviors has become an increasingly popular research area. The associations investigated in Panel C mainly focused on ethical behaviors; however, there were several other topics that found associations between uncertainty avoidance and other topics. Of the 14 articles in Panel C, seven of them were written in the past three years. Similar to the articles that listed uncertainty avoidance with Hofstede's other cultural dimensions, the associations between ethics and uncertainty avoidance is highly researched. Three of these articles focused on ethics [25] [92] [96], while four other articles researched similar topics such as corruption [115], software piracy [88] moral reasoning [45] and whistle-blowing [159]. The other articles examined associations between uncertainty avoidance and topics unrelated to ethics such as: brand loyalty [152], conformity [86], intolerance of deviations from group norms [148], religious factors [131], rules orientation [102], rule specific control systems [52] and technology skepticism [76].

In addition to sorting the Panel C articles based on topics, there are also differences in the suggested associations between uncertainty avoidance and other variables. Of the 14 articles that associate uncertainty avoidance a behavior, seven articles suggested positive relationships. In these studies uncertainty avoidance had a positive association with lower ethical problems [25], rules specific control systems [52], technology uncertainty [76], ethics [92], rules orientation [102], high brand loyalty [152] and intolerance of deviations from group norms [148]. In four articles, negative association was suggested between uncertainty avoidance and software piracy [88], ethical decision making [96], religious factors [131] and whistle-blowing in China [159]. The authors of the last three articles in Panel C suggested an association involving uncertainty avoidance; however, these articles did not indicate whether the associations were positive or negative. These articles discuss uncertainty avoidance having associations with moral reasoning [45], conformity [86] and corruption [115].

The three articles in Panel D of Table 2 are placed into a miscellaneous section as we were not able to categorize these articles into any of the prior groups. One article suggests uncertainty avoidance as a tool to compare firms [16]. Another discussed uncertainty avoidance as it relates to female expatriates [59] and the final article suggested using uncertainty avoidance in future research [66].

Literature Review Only Articles

Table 3 includes articles that used uncertainty avoidance in more in depth in a literature review. It separated into four panels. Panel A includes 13 articles that use uncertainty avoidance for definitions, hypothesis or one country studies. Panel B contains seven articles that utilize uncertainty avoidance in ethics studies. The 16 articles in Panel C use uncertainty avoidance in two counties studies. In the last subcategory, Panel D includes 12 articles that use uncertainty avoidance in studies that examine three or more countries.

TABLE 3

Literature Review Only Articles

Panel A: Definition, hypothesize and one country studies

Author(s)	Year	Theoretical Premise
Robertson et al.	2002	Describes UA
Rallapalli	1999	Describes UA
Vega	1997	Describes UA
Goodwin & Goodwin	1999	Describes UA and Hofstede's methodology in ranking countries
Parnell & Kedia	*1996*	*Describes UA and uses UA to explain Japan's success and negotiating effectiveness*
Thorne & Saunders	2002	Describes UA and some applications
Hoffman	1998	Describes UA and understanding cultural differences when expanding internationally

Yang et al.	2009	Describes UA and said UA has little impact on software piracy
Lin & Yeh	2009	Suggests UA positively associated with planning for future, working hard, and avoiding risky activities and decisions
Cohen et al.	1992	Suggests UA affects elements of accounting firms and behaviors
Volkema & Fleury	2002	Uses UA to describe US
Goel et al.	2005	Suggests UA associated with levels of trust
Woodbine	2004	Found no association between UA and Confucian work dynamism

Panel B: Use in studies of ethics and codes of ethics

Author(s)	Year	Theoretical Premise
De Bock & Van Kenhove	2010	Explained UA and suggested association with ethical beliefs
Armstrong	1996	Suggests association between UA and perceived ethical problems
Cherry et al.	2003	Suggests countries with high UA scores have fewer perceived ethical problems
Weaver	2001	Describes UA and suggest UA positively associated with focus on formal, visible codes of ethics and practices
Oumlil & Balloun	2009	Uses UA to explain external validity of Moroccan companies' codes of ethics
Salter et al.	2001	Associated likelihood of cheating to UA
Swaidan et al.	2006	Uses UA to hypothesize ethical differences and adherence ethical codes between developed and less developed countries

Panel C: Two country studies

Author(s)	Year	UA used to:
Tsui & Windsor	2001	Compare Australia and China
Srnka et al.	2007	Compare Austria and Turkey
Ge & Thomas	2008	Compare Canada and China
Fadil	1997	Compare Japanese and Hispanics
Van Es & Pels	2010	Compare Japan and Netherlands
Caldwell & Clapham	2003	Compare North Americans and Asians
Beekun et al.	2003	Compare US and Brazil
Tavakoli et al.	2003	Compare US and Croatia
Beekun et al.	2008	Compare US and Egypt
Sims & Keenan	1999	Compare US and Jamacia
Kim et al.	2010	Compare US and Korea
Daspro	2009	Compare US and Mexico
Vitell & Hidalgo	2006	Compare US and Spain
Lu et al.	1999	Compare US and Taiwan

Rawwas et al.	2005	Compare US and Turkey
Douglas & Wier	2005	Hypothesize differences between US and Chinese managers' incentives to create slack

Panel D: *Three plus country studies*

Author(s)	Year	UA used to compare:
Svensson et al.	1994	Australia, Canada, and Sweden; UA positively associated with formal rules about dealing with unpredictability
Nyaw & Ng	1994	Canada, Japan, Hong Kong and Taiwan
Husted	2002	Latin American countries
Gulbro & Herbig	*1998*	*Mexico, China, France, Japan, and Germany*
Sims & Gegez	2004	US, Australia, Israel, South Africa, and Turkey,
Bernardi & Guptill	2008	US, Canada, Colombia, Ecuador, Hong Kong, and Ireland
Baker & Veit	1998	US, Canada, Hong Kong, Japan, Singapore, and Thailand
Sower et al.	*1998*	*US, Canada, and Mexico*
Ramasamy & Yeung	2009	US, China, France, Germany, and Hong Kong
Shafer et al.	2007	US, Hong Kong, Malaysia, and Taiwan
Burnaz et al.	2009	US, Thailand and Turkey
Bernardi et al.	2009	Ethical decisions in US, Colombia, Ecuador, and South Africa
At-Twaijri & Al-Muhaiza	*1996*	*Gulf Cooperation Council Countries and finds a recent overall increase in UA due to political instability and threats of larger, neighboring countries*

Italicized data indicate publications in the International Journal of Value Based Management and Teaching Business Ethics.

The 13 articles in Panel A of Table 3 includes articles that are included in literature reviews for research involving uncertainty avoidance and can be subcategorized into three sections: definitions, hypothesis and one country studies. Seven of the articles in Panel A simply describe uncertainty avoidance [111] [104] [146] [49] [141] [54] [157]. Four of the articles use uncertainty avoidance to support research hypotheses. These articles suggest an association between uncertainty avoidance and Japan's success and negotiating effectiveness [94], planning for future [77], accounting firms and behaviors [33], trust levels [48] and Confucian work dynamism [156]. The last article in Panel A [149] used uncertainty avoidance to describe the United States.

Panel B of Table 3 contains seven articles that use uncertainty avoidance in ethics research. All articles suggested associations between uncertainty avoidance and specific ethical behaviors or actions. Five articles suggested associations between uncertainty avoidance and perceived ethical problems [3] [24], cheating [113], visible codes of ethics [153], and validity of codes of ethics in Morocco [93]. The other two articles in Panel B suggested associations between uncertainty avoidance and ethics without specifically mentioning any particular aspect of ethics [39] [136].

Fifteen of the 16 articles included in Panel C used uncertainty avoidance in the literature review to contrast two countries. The last article considered uncertainty avoidance when hypothesizing about behaviors in two countries. Panel C includes articles that compared Austria and Turkey [134], Japan and the Netherlands [145], Japanese and Hispanics [43], and North Americans and Asians [21]. Two articles include China in the two-country comparisons; Ge and Thomas [47] compared China and Canada, while Tsui and Windsor [143] compared China and Australia. Eight of the articles in Panel C compared the United States to: Brazil [12], Croatia [140], Egypt [11], Jamaica [124], Korea [67], Mexico [36], Spain [147], Taiwan [82] and Turkey [108]. The last article in Panel C used uncertainty avoidance to hypothesize differences between the United States and Chinese based on manager incentives [40].

The 12 articles in Panel D of Table 3 used uncertainty avoidance in studies that involved three or more countries; all of these articles use uncertainty avoidance to compare these countries. Eight of the twelve articles included the United States as one of the countries researched. These eight articles used uncertainty avoidance to compare the United States to Australia, Israel, South Africa and Turkey [123], Canada, Colombia, Ecuador, Hong Kong and Ireland [14] Canada, Hong Kong, Japan, Singapore and Thailand [9], Canada and Mexico [133], China, France, Germany and Hong Kong [105], Hong Kong, Malaysia and Taiwan [119], Thailand and Turkey [19], and Colombia, Ecuador and South Africa [15]. The remaining four articles, which did not include the United States, compared Australia, Canada and Sweden [135], Gulf Cooperation Council Countries [6], Mexico, China, France, Japan and Germany [51], and Latin American countries [58].

Alternative Cultural Index

As a simplified means to distinguish countries, Kogut and Singh [70] developed a calculation that uses Hofstede's cultural dimensions to construct a single score for each country rather than four individual scores (i.e., if one used Hofstede's dimensions directly). To compute their single index (Formula 1), Kogut and Singh calculate differences in cultures by subtracting each the target country's cultural dimensions (Hofstede) from the same cultural dimension of a control country, which is usually the United States. Once all of the cultural dimensions have been subtracted, the results are totaled for a single number score. Hofstede's cultural dimensions evaluate countries based on power distance, individualism, masculinity and uncertainty avoidance.

$$\text{Hofstede's Distance} = \left[\sum_{i=1}^{4} \{(I_{ii} - I_{iu}]^2\} / V_i\right] \Big/ 4 \qquad (1)$$

[70, p.422]

Where:

I_{ij}	Index for the *i*th cultural dimension and *j*th country
V_i	Variance of the index of the *i*th cultural dimension
u	Indicates the United States

Uhlenbruck's [144] cultural index (Formula 2) is the same as Kogut and Singh's index; however, instead of using Hofstede's cultural dimensions, Uhlenbruck used Trompenaar and Hampden-

Turner's dimensions. Uhlenbruck uses the same technique as Kogut and Singh, which results in single score for each country. Uhlenbruck uses three of Trompenaar and Hampden-Turner's cultural dimensions: individualism, specificity and attitude towards organizations. Table 4 contains three articles that use either the Kogut and Singh index or Uhlenbruck's index.

$$\text{Cultural Distance} = \left[\sum_{i \in Da}(I_{ia} - I_{it})^2 / V_i\right] \Big/ \left| D_{at} \right| \quad (2)$$

[144, p. 117]

Where:

I_{ia}	Index for the *i*th cultural dimension for the *a*th country
I_{it}	Index for the *i*th cultural dimension for the *t*th country
V_i	Variance of the index of the *i*th cultural dimension
$\left\| D_{at} \right\|$	Number of compared measures

Table 4 includes articles that utilize an alternative cultural dimensions evaluation index. The two alternative cultural dimension indices are Kogut and Singh and Uhlenbruck. Two articles use Kogut and Singh's Index to compare to uncertainty avoidance. Li [74] used Kogut and Singh to explain the likelihood of joint ventures, while Peng and Beamish [98], related Kogut and Singh to the predictability of general economic data. The third article used Uhlenbruck to explain cultural differences in ethics [50]. The final article used Kogut and Singh's index as well as Uhlenbruck. Voyer and Beamish [150] used both cultural indices to explain differences between Japan and other countries.

TABLE 4

Articles Employing a Cultural Index

Author(s)	Year	Sample	Theoretical Premise
Gopalan & Thomson	*2003*	*Theoritical*	*UA and Trompenaars were used to explain cultural differences in ethics.*
Li	2008	22,156 firms	UA and Kogut and Singh's Index used to explain likelihood of joint ventures.
Peng & Beamish	2008	50 countries	UA and Kogut and Singh's Index used to predict general economic data.
Voyer & Beamish	2004	9,546 investments 9 countries	Used Hofstede (Trompenaars and Hampden-Turner's) cultural dimensions and Kogut and Singh's Index (Uhlenbruck's Index) to explain differences between Japan and another country.

Italicized data indicate publications in the *International Journal of Value Based Management* and *Teaching Business Ethics*.

Using Uncertainty Avoidance as a Variable

Table 5 presents the 21 articles that use uncertainty avoidance as an independent variable in the research. As seen from this table, these hypotheses propose associations between cultures or behaviors. Similar to the previous tables, Table 5 is broken into three panels depending on article characteristics. Panel A presents the 9 articles that used survey data for statistical analysis. Panel B contains six articles that used country data in the analysis. Panel C has six articles that use alternative sources for the data.

Table 5

Articles Employing Uncertainty Avoidance as a Variable

Panel A: Human-participant studies

Author(s)	Year	Sample	Findings Relating to Culture
Lin & Ho	2010	466 adults 2 countries	UA higher in Taiwan than China
Auger et al.	2007	600 adults 6 countries	UA positively associated with worker, labor and individual rights
Kim & Kim	2010	125 adults South Korea	UA positively associated with good business, commitment, PR role and total corporate social responsibility
Tan & Chow	2009	58 adults 2 countries	Differences in attitudes found between US Caucasians and Chinese (in either China or US).
Peek et al.	2007	78 Students 3 countries	UA associated with attitudes concerning sexual harassment and whistle blowing among students from Canada, Mexico and US.
Arnold et al.	2007	294 adults 8 countries	UA not associated with measures of sensitivity.
Bernardi	2006	1,537 students 12 countries	UA positively associated with social desirability response bias.
Forsyth et al.	2008	30,230 adults 29 countries	UA positively associated with Relativism
Christie et al.	2003	345 adults 3 countries	UA associated with questionable ethical behavior in US, Korea and India

Panel B: Country-based studies

Author(s)	Year	Sample	Findings Relating to uncertainty avoidance
Baughn et al.	2010	125 countries	UA positively associated with OECD convention ratification
Li et al.	2008	133 countries	UA positively associated with Corruption Perception Index
Davis & Ruhe	2003	42 countries	UA positively associated with preference for bureaucratic structures, nationalism, Corruption Perception Index, and unethical behavior.

Sanyal	2005	30 countries	UA positively associated with Corruption Perception Index
Peng & Lin	2009	51 countries	UA negatively associated with individualism
Husted	2000	50 countries	No association between UA and other cultural dimensions with software piracy using World Bank data.

Panel C: Other studies

Author(s)	Year	Sample	Findings Relating to Culture
Johan & Najar	2010	123 Firms	UA positively associated with high efficiency and high performance fees
Cummings et al.	2010	528 Firms	UA positively (negatively) associated with the Masculinity index and Corruption Perception index (Power Distance index)
Clements et al.	2009	104 Firms	UA positively (negatively) associated with PDI (IDV)
Scholtens & Dam	2007	2683 Firms 24 countries	UA is positively associated with a firm's ethical policies
Alas	2006	Estonian Organizations	UA positively (negatively) associated with values (practices) of ethics
Chen et al.	*2008*	*2,018 firms 55 countries*	*UA not significant for variables examined*

Italicized data indicate publications in the *International Journal of Value Based Management* and *Teaching Business Ethics*.

The Panel A of Table 5 includes data from surveys of either students or adult college graduates in the workforce. Panel A contains two articles that use uncertainty avoidance to suggest differences between cultures. While Lin & Ho [80] contrast the differences in uncertainty avoidance between Taiwan and China, Tan & Chow [137] explain cultural differences between American Caucasians and people of Chinese decent who reside in either China or the United States.

With the exception of the previously mentioned articles, Panel A organizes articles that illustrate associations between uncertainty avoidance and individual behaviors. Four articles describe potential positive associations between uncertainty avoidance and worker, labor and individual rights [7]; good business, commitment, public relations role and corporate social responsibility [68] social desirability response bias [14] and, relativism [46]. In addition to the positive associations that were found with uncertainty avoidance, there was also one negative relationship found. Arnold et al. [4] discovered a negative association between uncertainty avoidance and measures of sensitivity.

The last two articles included in Panel A that found a relationship between uncertainty avoidance and a behavior; however, it was not stated whether this relationship was positive or negative. Peek et al. [97] associated whistle-blowing and sexual harassment to uncertainty avoidance; Christie et al. [29] associated questionable ethical behavior to uncertainty avoidance.

The second section, Panel B contains articles that utilize uncertainty avoidance as an independent variable in the research and publically available database information from a variety of sources in the analysis. The sample sizes range from 30-to-133 countries. Panel B has six articles; four found positive associations with uncertainty avoidance, one found a negative association and one found no association. Uncertainty avoidance was positively associated with OECD convention ratification [10], corruption perception index [75][37][116] and preference for bureaucratic structures, nationalism and unethical behavior [37]. Uncertainty avoidance was negatively associated with individualism [99]. Husted [57] found no association between uncertainty avoidance with software piracy.

The final section of Table 5 has articles with firm-level sources. Panel C has six articles that used alternative sample sources. Johan and Najar [64] collected a sample from 123 firms which represented 23 countries. This article found that uncertainty avoidance is positively associated with high efficiency and high performance fees. The next article included a sample from 528 firms in 20 countries [35]. This article described that uncertainty avoidance is positively associated with Masculinity Index and the Corruption Perception Index as well as a negatively associated with the Power Distance Index. Clements et al. [30] had a sample of 104 accounting firms; this article found that uncertainty avoidance is positively associated with power distance and is negatively associated with individualism. Scholtens and Dam [117] collected data from 2683 firms that came from 24 countries; these authors showed that uncertainty avoidance is positively related with a firm's ethical policies. The third article in Panel C had a sample that utilized information from Estonian organizations [1]. While Alas found that uncertainty avoidance is positively associated with values of ethics, uncertainty avoidance was negatively associated with the practice of ethics. The last article [23] examined in Panel C had a sample from 2,018 firms representing 55 countries. This article found no significant associations between uncertainty avoidance and any of the variables examined.

CONCLUSIONS

This study examined the extent and nature of the use of Hofstede's uncertainty avoidance in articles published in the *Journal of Business Ethics,* the *International Journal of Value Based Management* and *Teaching Business Ethics* for a period of 29 years from 1982 to 2010. The beginning of this timeframe (1982) with the first issue of the *Journal of Business Ethics* is two years after Hofstede's (1980) initial publication of his four cultural dimensions. While these dimensions were introduced in 1980, our data indicate that the first article to use uncertainty avoidance in each of these three journals was: 1992 for the *Journal of Business Ethics,* 1990 for the *International Journal of Value Based Management* and 1998 for *Teaching Business Ethics*.

Our data also indicate that 83.8 % (129 of the 154) of the articles citing Hofstede's work used the uncertainty avoidance construct as part of their literature reviews to make the point about differences in cultures, which is significantly above the 50.0% noted by Rapp et al. [106] in the *Journal of International Business Studies*. However, only four of the 154 articles (2.6 %) used Hofstede's cultural constructs in the computation of cultural distance indices; this compares to 29 articles of the 118 articles (24.6 %) that Rapp et al. [106] reported that used this methodology. This difference suggests that the use of Kogut and Singh's [70] index has not been adopted by

ethics scholars, which presents an avenue for future research. The remaining 21 articles (13.6%) were the only articles that used uncertainty avoidance as an independent or control variable. It appears that the growing influence of Hofstede's uncertainty avoidance construct on international ethics research (Figure 1) is primarily as a basis in the literature review for the hypotheses development.

A limitation of our research is that it only considered the articles in three ethics journals: the *Journal of Business Ethics,* the *International Journal of Value Based Management* and *Teaching Business Ethics*; of these journals, the last two have not been published since 2003. This limitation also provides the opportunity for future research; future research could examine other research-specific journals to determine whether the results of this research are supported.

REFERENCES

[1]Alas, R. *Ethics in countries with different cultural dimensions*, Journal of Business Ethics, 2006,69 **(**3), 237-247.

[2]Al-Khatib, J. A., Rawwas, M. Y. A. and Vitell, S. J., *Organizational Ethics in Developing Countries: A Comparative Analysis*, Journal of Business Ethics, 2004,55 (4), 309-322.

[3]Armstrong, R. W., *The relationship between culture and perception of ethical problems in international Marketing*, Journal of Business Ethics, 1996,15 (11), 1199-1208.

[4]Arnold, D. F., R. Bernardi, P. E. Neidermeyer and Schmee, J., *The Effect of Country and Culture on Perceptions of Appropriate Ethical Actions Prescribed by Codes of Conduct: A Western European Perspective among Accountants,* Journal of Business Ethics, 200,**770** (4), 327-340.

[5]Atakan, M. G. S. Burnaz., S and Topcu, Y. I.., *An Empirical Investigation of the Ethical Perceptions of Future Managers with a Special Emphasis on Gender – Turkish Case*, Journal of Business Ethics, 2008,82 (3), 573-586.

[6]At-Twaijri, M. I. and Al-Muhaiza, I. A. *Hofstede's Cultural Dimensions in the GCC Countries: An Empirical Investigation,* International Journal of Value-Based Management 1996,9 (2), 121-131.

[7]Auger, P., Devinney, T. M. and Louviere, J. J., *Using Best-Worst Scaling Methodology to Investigate Consumer Ethical Beliefs Across Countries,* Journal of Business Ethics, 2007,70 (3), 299-326.

[8]Aygün, Z. K., Arslan,M. and Güney, S.. *Work Values of Turkish and American University Students,* Journal of Business Ethics, 2008,80 (2), 205-223.

[9]Baker, H. K. and Veit, E. T. *A Comparison of Ethics of Investment Professionals: North America versus Pacific Rim Nations*, Journal of Business Ethics, 1998,17 (8), 917-937.

[10]Baughn, C., Bodie, N. L., Buchanan, M. A. and Bixby, M. B. *Bribery in International Business Transactions*, Journal of Business Ethics, 2010,92**,** (1), 15-32.

[11]Beekun, R. I., Hamdy, R., Westerman, J. W. and HassabElnaby, H. R.., *An Exploration of Ethical Decision-making Processes in the United States and Egypt*, Journal of Business Ethics, 2008, 82 (3), 587-605.

[12]Beekun, R. I.,. Stedham, Y and Yamamura, J. H., *Business Ethics in Brazil and the U.S. A Comparative Investigation,* Journal of Business Ethics **,**2003,42 (3), 267-279.

[13]Beets, S. D., *Understanding the Demand-Side Issues of International Corruption, Journal of Business Ethics,* 2005,57 (1), 65-81.

[14]Bernardi, R. A., *Associations between Hofstede's Cultural Constructs and Social Desirability Response Bias*, Journal of Business Ethics, 2006,65 (1), 46-56.
[14]Bernardi, R. and Guptill, S., *Social Desirability Response Bias, Gender, and Factors Influencing Organizational Commitment: An International Study*, Journal of Business Ethics, 2008,81 (4), 797-809.
[15]Bernardi, R., Witek, M. B. and Melton, M. R., *A Four-Country Study of the Associations between Bribery and Unethical Actions*, Journal of Business Ethics, 2009,84 (3), 389-403.
[16]Bowen, F., *Corporate Social Strategy: Competing Views from Two Theories of the Firm*, Journal of Business Ethics, 2007,75 (1), 97-113.
[17]Brinkmann, J., *Business and Marketing Ethics as Professional ethics, Concepts, Approaches and Typologies*, Journal of Business Ethics, 2002,41 (1), 159-177.
[18]Brody, R. G., Coulter, J. M. and Lin. S., *The Effect of National Culture on Whistle-Blowing Perceptions*, Teaching Business Ethics, 1999,3 (4), 385-400.
[19]Burnaz, S.,. Atakan, M. G. S, Topcu, Y. I. and Singhapakdi, A., *An Exploratory Cross-Cultural Analysis of Marketing Ethics: The Case of Turkish, Thai, and American Businesspeople*, Journal of Business Ethics, 2009, 90 (3), 371-382.
[20]Burton, B. K., Farh, J. and Hegarty, W. H. , *A Cross-Cultural Comparison of Corporate Social Responsibility Orientation: Hong Kong vs. United States Students*, Teaching Business Ethics, 2000,4 (2), 151-167.
[21]Caldwell C. and Clapham, S. E., *Organizational Trustworthiness: An International Perspective*, Journal of Business Ethics, 2003, 47 (4), 349-364.
[22]Chandler, J. D and Graham, J. L., *Relationship-Oriented Cultures, Corruption, and International Marketing Success*, Journal of Business Ethics, 2010, 92 (2), 251-267.
[23]Chen, Y., Yaşar, M. and Rejesus, R. M., *Factors Influencing the Incidence of Bribery Payouts by Firms: A Cross-Country Analysis*, Journal of Business Ethics, 2008, 77 (2), 231-244.
[24]Cherry, J., Lee, M. and Chien, C. S., *A Cross-Cultural Application of a Theoretical Model of Business Ethics: Bridging the Gap between Theory and Data*, Journal of Business Ethics, 2003, 44 (4), 359-376.
[25]Cherry, J., *The Impact of Normative Influence and Locus of Control on Ethical Judgments and Intentions: a Cross-Cultural Comparison*, Journal of Business Ethics, 2006,68 (2), 113-132.
[26]Cheung, C. and Chan, A. C., *Philosophical foundations of Eminent Hong Kong Chinese CEOs' Leadership*, Journal of Business Ethics, 2005, 60 (1), 47-62.
[27]Chhokar, J. S., Zhuplev, A., Fok, L. Y. and Hartman S. J., *The Impact of Culture on Equity Sensitivity Perceptions and Organizational Citizenship Behavior: A Five-Country Study*, International Journal of Value-Based Management, 2001,14 (1), 79-98.
[28]Chiu, H. C., Hsieh, Y. C. and Wang, M. C., *How to Encourage Customers to Use Legal Software*, Journal of Business Ethics, 2008,80 (3), 583-595.
[29]Christie, P. M. J., Kwon, I. G., Stoeberl, P. A. and Baumhart, R., *A Cross-Cultural Comparison of Ethical Attitudes of Business Managers: India Korea and the United States*, Journal of Business Ethics, 2003,46 (3), 263-287.
[30]Clements, C. E., Neill J. D. and Stovall, O. S., *The Impact of Cultural Differences on the Convergence of International Accounting Codes of Ethics*, Journal of Business Ethics, 2009, 90 (3), 383-391.
[31]Clikeman, P. M., M. A. Geiger and B. T. O'Connell. 2001, 'Student Perceptions of Earnings Management: The Effects of National Origin and Gender', *Teaching Business Ethics* **5** (4), 389-410.

[32]Cohen, J. R. and Single, L. E., *An Examination of the Perceived Impact of Flexible Work Arrangements on Professional Opportunities in Public Accounting*, Journal of Business Ethics, 2001,32 (4), 317-328.
[33]Cohen, J. R., Pant, L.W. and Sharp, D. J., *Cultural and Socioeconomic Constraints on International Codes of Ethics: Lessons from Accounting*, Journal of Business Ethics, 1992,11 (9), 687-700.
[34]Collins, J. D., Uhlenbruck, K. and Rodriguez, P., *Why Firms Engage in Corruption: A Top Management Perspective*, Journal of Business Ethics, 2009,87 (1), 89-108.
[35]Cummings, D., Fleming, G. Johan, S. and Takeuchi, M., *Legal protection, corruption and private equity returns in Asia*, Journal of Business Ethics, 2010, 95 (2), 173-193.
[36]Daspro, E., *An Analysis of U.S. Multinationals' Recruitment Practices in Mexico*, Journal of Business Ethics ,2009,87 (1), 221-232.
[37]Davis, J. and Ruhe, J. A., *Perceptions of Country Corruption: Antecedents and Outcomes*, Journal of Business Ethics, 2003,43 (4), 277-288.
[38]Davis, M. A, Johnson, N. B. and Ohmer, D. G., *Issue-Contingent Effects on Ethical Decision Making: A Cross-Cultural Comparison*, Journal of Business Ethics, 1998,17 (4), 373-389.
[39]De Bock, T. and Van Kenhove, P., *Consumer Ethics: The Role of Self-Regulatory Focus*, Journal of Business Ethics, 2010,97 (2), 241-255.
[40]Douglas, P. C. and Wier. B, *Cultural and Ethical Effects in Budgeting Systems: A Comparison of U.S. and Chinese Managers*, Journal of Business Ethics, 2005, 60 (2), 159-174.
[41]Dunn, P and Shome, A., *Cultural Crossvergence and Social Desirability Bias: Ethical Evaluations by Chinese and Canadian Business Students*, Journal of Business Ethics, 2009,85 (4), 527-543.
[42]Erben, G. S. and Güneşer, A. B., *The Relationship between Paternalistic Leadership and Organizational Commitment: Investigating the Role of Climate Regarding Ethics*, Journal of Business Ethics, 2008,82 (4), 955-968.
[43]Fadil, P. A., *The Influence of Cultural Adaptation on Hispanic-Americans and Japanese-Americans*, International Journal of Value-Based Management, 1997,10 (1), 55-67.
[44]Fam, K. S. and Waller, D. S., *Advertising Controversial Products in the Asia Pacific: What Makes Them Offensive?*, Journal of Business Ethics, 2003,48 (3), 237-259.
[45]Fleming, D. M., Chow, C. W. and Su, W., *An Exploratory Study of Chinese Accounting Students' and Auditors' Audit-specific Ethical Reasoning*, Journal of Business Ethics, 2010,94 (3), 353-369.
[46]Forsyth, D. R., O'Boyle, E. H. and McDaniel, M. A., *East Meets West: A Meta-Analytic Investigation of Cultural Variations in Idealism and Relativism*, Journal of Business Ethics, 2008,83 (4), 813-833.
[47]Ge, L. and Thomas, S., *A Cross-Cultural Comparison of the Deliberative Reasoning of Canadian and Chinese Accounting Students*, Journal of Business Ethics, 2008,82 (1), 189-211.
[48]Goel, S., Bell, G. G. and Pierce, J. L., *The Perils of Pollyanna: Development of the Over-Trust Construct*, Journal of Business Ethics, 2005,58 (1), 203-218.
[49]Goodwin J. and Goodwin, D., *Ethical Judgment across Cultures: A Comparison between Business Students from Malaysia and New Zealand*, Journal of Business Ethics, 1999,18 (3), 267-281.
[50]Gopalan, S. and Thomson, N., *National Cultures, Information Search Behaviors and the Attribution Process of Cross-National Managers: A Conceptual Framework*, Teaching Business Ethics, 2003,7 (3), 313-328.

[51]Gulbro, R. D. and Herbig, P., *Cultural Differences in International Negotiating*, International Journal of Value-Based Management, 1998,11 (3), 265-273.
[52]Haswell, S., Jubb, P. and Wearing, B., *Accounting Students and Cheating: A Comparative Study for Australia, South Africa and the UK*, Teaching Business Ethics, 1999,3 (3), 211-239
[53]Hay, D., Larres, P. M., Oyelere, P. and Fisher, A., *The Ethical Perception of Undergraduate Students in Computer-Related Situations: An Analysis of the Effects of Culture, Gender and Prior Education*, Teaching Business Ethics, 2001,5 (3), 331-356.
[54]Hoffman, J. J., *Evaluating International Ethical Climate: A Goal Programming Model*, Journal of Business Ethics, 1998,17 (16), 1861-1869.
[55]Hofstede, G., *Organizations and cultures: Software of the mind*, New York: McGraw-Hill,1991
[56]Hofstede, G., *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations across Countries*. Beverly Hills, CA: Sage Publications. 1980.
[57]Husted, B., *The Impact of National Culture on Software Piracy*, Journal of Business Ethics, 2000,26 (3), 197-211.
[58]Husted, B., *Culture and International Anti-Corruption Agreements in Latin America*, Journal of Business Ethics ,2002,37 (4), 413-422.
[59]Insch, G. S., McIntyre, N. and Napier, N. K., *The Expatriate Glass Ceiling: The Second Layer of Glass*, Journal of Business Ethics, 2008,83 (1), 19-28.
[60]Jin, K. G. and Drozdenko, R. G., *Relationships among Perceived Organizational Core Values, Corporate Social Responsibility, Ethics, and Organizational Performance Outcomes: An Empirical Study of Information Technology Professionals*, Journal of Business Ethics, 2010,92 (3), 341-359.
[61]Jin, K. G., Drozdenko, R. and Bassett, R., *Information Technology Professionals' Perceived Organizational Values and Managerial Ethics: An Empirical Study*, Journal of Business Ethics 2007,71 (2), 149-159.
[62]Jing, R. and Graham, J. L., *Values Versus Regulations: How Culture Plays Its Role*, Journal of Business Ethics *2008,80 (4), 791-806.*
[63]Jones, M. T., *The Institutional Determinants of Social Responsibility*, Journal of Business Ethics **,**1999,20 (2), 163-179.
[64]Johan S. A. and Najar, D., *The role of corruption, culture, and law in investment fund manager fees*, Journal of Business Ethics, 2010,95 (2), 147-172.
[65]Jose, A. and Thibodeaux, M. S., *Institutionalization of Ethics: The Perspective of Managers*, Journal of Business Ethics, 1999,22 (2), 133-143.
[66]Khatri, N. and Tsang, E. W. K., *Antecedents and Consequences of Cronyism in Organizations*, Journal of Business Ethics, 2003,43 (4), 2889-303.
[67]Kim, G. S., Lee, G. Y. and Park, K., *A Cross-National Investigation on How Ethical Consumers Build Loyalty toward Fair Trade Brands*, Journal of Business Ethics, 2010,96 (4) 589-611.
[68]Kim, Y. and Kim, S. Y., *The Influence of Cultural Values on Perceptions of Corporate Social Responsibility: Application of Hofstede's Dimensions to Korean Public Relations Practitioners*, Journal of Business Ethics, 2010,91 (4), 485-500.
[69]Kim, Y., *Ethical Standards and Ideology among Korean Public Relations Practitioners*, Journal of Business Ethics, 2003,42 (3), 209-223.
[70]Kogut, B. and Singh, H., *The Effect of National Culture on the Choice of Entry Mode*, Journal of International Business Studies, 1988,19 (3), 411-432.

[71]Lämsä, A., Säkkinen, A. and Turjanmaa, P., *Values and Their Change during the Business Education – A Gender Perspective*, International Journal of Value-Based Management, 2000,13 (3), 203-213.
[72]Leong, S. M., Tan, H. H. and Loh, M. S., *When the Cat's Away: A Content Analysis of MNC Overseas Recruitment Print Ads*, Journal of Business Ethics, 2004,49 (2), 115-127.
[73]Leung, A. S. M., Liu, X. and Liu, S., *Moral Schemas and Business Practices: The Ethics of Guangzhou Migrant Marketers*, Journal of Business Ethics, 2009,88 (1), 11-23.
[74]Li, J., Moy, J., Lan, K. and Chu, W. L., *Institutional Pillars and Corruption at the Societal Level*, Journal of Business Ethics, 2008,83 (2), 327-229.
[75]Li, N., *Religion, Opportunism, and International Market Entry via Non-Equity Alliance or Joint Ventures*, Journal of Business Ethics, 2008, 80 (4), 771-789.
[76]Li, S., Karande, K. and Zhou, D., *The Effect of the Governance Environment on Marketing Channel Behaviors: The Diamond Industries in the U.S., China, and Hong Kong*, Journal of Business Ethics, 2009, 88 (3), 453-471.
[77]Lin, C. L. and Yeh, J. T., *Comparing Society's Awareness of Women: Media-Portrayed Idealized Images and Physical attractiveness*, Journal of Business Ethics, 2009,90 (1), 61-79.
[78]Lin, C. L., *Sexual Issues: The Analysis of Female Role Portrayal Preferences in Taiwanese Print Ads*, Journal of Business Ethics, 2008, 83 (3), 409-418.
[79]Lin, C. P., *Modeling Corporate Citizenship, Organizational Trust, and Work Engagement Based on Attachment Theory*, Journal of Business Ethics, 2010, 94 (4), 517-531.
[80]Lin, L. H. and Ho, Y. L., *Guanxi and OCB: The Chinese Cases*, Journal of Business Ethics, 2010, 96 (2), 285-298.
[81]Lowe, S., *Codes of Conduct in Hong Kong Organizations: A Preliminary Analysis*, International Journal of Value-Based Management, 1996, 9 (3), 211-225.
[82]Lu, L. C., Rose, G. M. and Blodgett, J. G., *The Effects of Cultural Dimensions on Ethical Decision Making in Marketing: An Exploratory Study*, Journal of Business Ethics, 1999, 18 (1), 91-105.
[83]MacNab, B. R. and Worthley, R., *Self-Efficacy as an Intrapersonal Predictor for Internal Whistleblowing: A US and Canada Examination*, Journal of Business Ethics, 2008, 79 (4), 407-421.
[84]Marta, J. K. M., Attia, A. Singhapakdi, A. and Atteya, N., *A Comparison of Ethical Perceptions and Moral Philosophies of American and Egyptian Business Students*, Teaching Business Ethics**,** 2003, 7 (1), 1-20.
[85]Marta, J., Heiss, C. M. and. De Lurgio, S. A., *An Exploratory Comparison of Ethical Perceptions of Mexican and U.S. Marketers*, Journal of Business Ethics, 2008, 82 (3), 539-555.
[86]Mills, G. R., Austin, S. A., Thomson D. S. and Devine-Wright, H., *Applying a Universal Content and Structure of Values in Construction Management*, Journal of Business Ethics, 2009,90 (4), 473-501.
[87]Monga, M., *'Managers' Moral Reasoning: Evidence from Large Indian Manufacturing Organizations*, Journal of Business Ethics, 2007, 71 (2), 179-194.
[88]Moores, T., *An Analysis of the Impact of Economic Wealth and National Culture on the Rise and Fall of Software Piracy Rates*, Journal of Business Ethics, 2008, 81 (1), 39-51.
[89]Nebenzahl, I. D., Jaffe, E. D. and Kavak, B., *Consumers' Punishment and Rewarding Process via Purchasing Behavior*, Teaching Business Ethics, 2001, 5 (3), 283-305.
[90]Ng, E. S. and Burke, R. J., *Predictor of Business Students' Attitudes toward Sustainable Business Practices*, Journal of Business Ethics, 2010, 95 (4), 603-615.

[91]Nyaw, M. and Ng, I., *A Comparative Analysis of Ethical Beliefs: A Four Country Study*, Journal of Business Ethics, 1994, 13 (7), 543-555
[92]O'Higgins, E. and Kelleher, B., *Comparative Perspectives on the Ethical Orientations of Human Resources, Marketing and Finance Functional Managers*, Journal of Business Ethics 2005, 56 (3), 275-288.
[93]Oumlil, A. B. and Balloun, J. L., *Ethical Decision-Making Differences between American and Moroccan Managers*, Journal of Business Ethics, 2009, 84 (4), 457-478.
[94]Parnell, J. A. and Kedia, B. L., The Impact of National Culture on Negotiating Behaviors Across Borders, *International Journal of Value-Based Management*, 1996, 9 (1), 45-61.
[95]Parnell, J. A. and Hatem, T., *The Cultural Specificity of Management Constructs: An Empirical Examination*, International Journal of Value-Based Management, 1997, 10 (3), 247-271.
[96]Patel, T. and Schaefer, A., *Making Sense of the Diversity of Ethical Decision Making in Business: An Illustration of the Indian Context*, Journal of Business Ethics, 2009, 90 (2), 171-186.
[97]Peek, L., Roxas, M., Peek, G., Robinchaud, Y. and Covarrubias Salazar, B. E., et al., *NAFTA Students' Whistle-Blowing Perceptions: A Case of Sexual Harassment*, Journal of Business Ethics, 2007,74 (3), 219-231.
[98]Peng, G. Z. and Beamish, P. W., *The Effect of National Corporate Responsibility Environment on Japanese Foreign Direct Investment*, Journal of Business Ethics, 2008, 80 (4), 677-695.
[99]Peng, Y. S. and Lin, S. S., *National Culture, Economic Development, Population Growth and Environmental Performance: The Mediating Role of Education*, Journal of Business Ethics, 2009, 90 (2), 203-219.
[100]Peterson, R. A., Albaum, G. Merunka, D. Munuera J. L. and Smith, S. M.., *Effects of Nationality, Gender, and Religiosity on Business-Related Ethicality*, Journal of Business Ethics, 2010, 96 (4), 573-587.
[101]Phau, I. and Kea, G., *Attitudes of University Students toward Business Ethics: A Cross-National Investigation of Australia, Singapore and Hong Kong*, Journal of Business Ethics, 2007, 72 (1), 61-75.
[102]Preble, J. F. and Hoffman, R. C., *The Nature of Ethics Codes in Franchise Associations around the Globe*, Journal of Business Ethics, 1999, 18 (3), 239-253.
[103]Primeaux, P. S. M., Karri, R. and Coldwell, C., *Cultural Insights to Justice: A Theoretical Perspective through a Subjective Lens*, Journal of Business Ethics, 2003, 46 (2), 187-199.
[104]Rallapalli, K. C., *A Paradigm for Development and Promulgation of Global Code of Marketing Ethics*, Journal of Business Ethics, 1999, 18 (1), 125-137.
[105]Ramasamy, B. and Yeung, M., *Chinese Consumers' Perception of Corporate Social Responsibility (CSR)*, Journal of Business Ethics, 2009, 88 (1), 119-132.
[106]Rapp, J. K., Bernardi, R. A., and Bosco, S. M., *Examining the Use of Hofstede's Uncertainty Avoidance Construct in International Research: A 25-Year Review*, International Business Research Journal, 2011, 4 (1), 3-15.
[107]Rawwas, M. Y. A., Patzer, G. L. and Vitell, S. J., *A Cross-Cultural Investigation of the Ethical Values of Consumers: The Potential Effect of War and Civil Disruption*, Journal of Business Ethics, 1998, 17 (4), 373-389.

[108]Rawwas, M. Y. A., Swaidan, Z. and Oyman, M., *Consumer Ethics: A Cross-Cultural Study of the Ethical Beliefs of Turkish and American Consumers*, Journal of Business Ethics, 2005, 57 (2), 183-195.
][109]Resick, C. J., Hanges, P. J., Dickson M. W and Mitchelson, J. K., *A Cross-Cultural Examination of the Endorsement of Ethical Leadership*, Journal of Business Ethics**,** 2006, 63 (4), 345-359.
[110]Robertson C. and Fadil, P. A., *Ethical Decision Making in Multinational Organization: A Culture-Based Model*, Journal of Business Ethics, 1999, 19 (4), 385-392.
[111]Robertson, C. J., Crittenden, W. F., Brady, M K. and Hoffman, J. J., *Situational Ethics across Borders: A Multicultural Examination*, Journal of Business Ethics, 2002, 38 (4), 327-338.
[112]Robertson, D. C. and Schlegelmilch, B. B., *Corporate Institutionalization of Ethics in the United States and Great Britain*, Journal of Business Ethics, 1993, 12 (4), 301-312.
[113]Salter, S. B., Guffey, D. M. and McMillan, J. J., *Truth, Consequences and Culture: A Comparative Examination of Cheating and Attitude about Cheating among U.S. and U.K. Students*, Journal of Business Ethics, 2001, 31 (1), 37-50.
[114]Sama, L. M. and Papamarcos, S. D., *Hofstede's I-C Dimension as Predictive of Allocative Behaviors: A Meta-Analysis*, International Journal of Value-Based Management, 2000, 13 (2), 173-188.
[115]Sanchez, J. I., Gomez, C. and Wated, G., *A Value-based Framework for Understanding Managerial Tolerance of Bribery in Latin America*, Journal of Business Ethics, 2008, 83 (2), 341-352.
[116]Sanyal, R., *Determinants of Bribery in International Business: The Cultural and Economic Factors*, Journal of Business Ethics, 2005, 59 (1), 139-145.
[117]Scholtens, B. and Dam, L., *Cultural Values and International Differences in Business Ethics*, Journal of Business Ethics, 2007, 75 (3), 273-284.
[118]Schwartz, M. S., *Corporate Efforts to Tackle Corruption: An Impossible Task?" The Contribution of Thomas Dunfee*, Journal of Business Ethics, 2009, 88 (4), 823-832.
[119]Shafer, W. E., Fukukawa, K., and Lee, G. M., *Values and the Perceived Importance of Ethics and Social Responsibility: The U.S. versus China*, Journal of Business Ethics, 2007, 70 (3), 265-284.
[120]Sidani, Y. M. and Jamali, D., *The Egyptian Worker: Work Beliefs and Attitudes*, Journal of Business Ethics, 2010, 92 (3), 433-450.
[121]Sidani, Y. M. and Thornberry, J., *The Current Arab Work Ethic: Antecedents, Implications, and Potential Remedies*, Journal of Business Ethics, 2010, 91 (1), 35-49.
[122]Simga-Mugan, C., Daly, B. A., Onkal D., and Kavut L., *The Influence of Nationality and Gender on ethical Sensitivity: An Application of the Issue-Contingent Model*, Journal of Business Ethics, 2005, 57 (2), 139-159.
[123]Sims, R. L. and Gege, A. E., *Attitudes towards Business Ethics: A Five Nation Comparative Study*, Journal of Business Ethics, 2004, 50 (3), 253-265.
[124]Sims, R. L. and Keenan, J. P., *A Cross-Cultural Comparison of Managers' Whistleblowing Tendencies*, International Journal of Value-Based Management, 1999, 12 (2), 137-151.
[125]Sims, R. L., *Support for the Use of Deception within the Work Environment: A Comparison of Israeli and United States Employee Attitudes*, Journal of Business Ethics, 2002, 35 (1), 27-34.
[126]Singhapakdi, A. and Vitell, S. J., *From the Guest Editors International Marketing Editors*, Journal of Business Ethics, 1999, 18 (1), 1-2.

[127]Singhapakdi, A., Marta, J. K. M. Rao, C. P. and Cicic, M., *Is Cross-Cultural Similarity an Indicator of Similar Marketing Ethics?*, Journal of Business Ethics, 2001, 32 (1), 55-68.
[128]Singhapakdi, A., Gopinath, M. Marta J. K. and Carter, L. L., *Antecedents and Consequences of Perceived Importance of Ethics in Marketing Situations: A Study of Thai Businesspeople*, Journal of Business Ethics, 2008, 81 (4), 887-904.
[129]Sison, A. J.G., *The Cultural Dimension of Codes of Corporate Governance: A Focus on the Olivencia Report*, Journal of Business Ethics, 2000, 27 (1), 181-192.
[130]Smith, A. and Hume, E. C., *Linking Culture and Ethics: A Comparison of Accountants' Ethical Belief Systems in the Individualism/Collectivism and Power Distance Contexts*, Journal of Business Ethics, 2005, 62 (30), 209-220.
[131]Sood, J., *A Cross-Cultural Study of Value Emphasized in Education*, International Journal of Value-Based Management, 1998, 11 (3), 215-223.
[132]Sotto, F. C. and Kohls, J. J., *White-Collar Work Values and Attitudes: A Cross-Cultural Comparison of the U.S. and the Philippines*, International Journal of Value-Based Management, 1990, 3 (2), 21-35
[133]Sower, V. E., Abshire, R. D. and Shankman, N. A.., *A Cross-Cultural Study of the Ethical Orientations of Senior-Level Business Students*, Teaching Business Ethics, 1998, 1 (4), 379-397
[134]Srnka, K. J., Gegez, A. E. and Arzova, S. B., *Why Is It (Un-)ethical? Comparing Potential European Partners: A Western Christian and An Eastern Islamic Country – On Arguments Used in Explaining Ethical Judgments*, Journal of Business Ethics, 2007,74 (2), 101-118.
[135]Svensson, G., Wood, G., Singh, J., Carasco, E. and Callaghan, M., *Ethical Structures and Processes of Corporations Operating in Australia, Canada, and Sweden: A Longitudinal and Cross-Cultural Study*, Journal of Business Ethics, 2009, 86 (4), 485-506.
[136]Swaidan, Z., Vitell,S. J., Rose, G. M. and Gilbert, F. W., *Consumer Ethics: The Role of Acculturation in U.S. Immigrant Populations*, Journal of Business Ethics, 2006, 64 (1), 1-16.
[137]Tan, J. and Chow, I. H., *Isolating Cultural and National Influence on Value and Ethics: A Test of Competing Hypotheses*, Journal of Business Ethics, 2009, 88 (1), 197-210.
[138]Tan, J., *Breaking the 'Bamboo Curtain' and the 'Glass Ceiling': The Experience of Women Entrepreneurs in High-Tech Industries in an Emerging Market*, Journal of Business Ethics, 2008, 80 (3), 547-564.
[139]Tang, L., *An Integral Model of Collective Action in Organizations and Beyond*, Journal of Business Ethics, 2008,80 (2), 249-261.
[140]Tavakoli, A. A., Keenan, J. P. and Cranjak-Karanovic, B., *Culture and Whistle-Blowing: An Empirical Study of Croatian and United States Managers Utilizing Hofstede's Cultural Dimensions*, Journal of Business Ethics, 2003, 43 (1), 49-64.
[141]Thorne, L. and Saunders, S. B., *The Socio-Cultural Embeddedness of Individuals' Ethical Reasoning in Organizations (Cross-cultural Ethics)*, Journal of Business Ethics, 2002,35 (1), 1-14.
[142]Transparency International: 2010. *Corruption Perceptions Index*, accessed on Sept. 18, 2011, http://www.transparency.org/policy_research/surveys_indices/cpi/2010.
[143]Tsui, J. and Windsor, C., *Some Cross-Cultural Evidence on Ethical Reasoning*, Journal of Business Ethics, 2001, 31 (2), 143-150.
[144]Uhlenbruck, K., *Developing Acquired Foreign Subsidiaries: The Experience of MNES in Transition Economies*, Journal of International Business Studies, 2004, 35 (2), 109-123.

[145]Van Es, R. and Pels, T., *The Political Dimension: Added Value for Cross-Cultural Analysis. Nozawa and Smits, Two CEOs and Their Public Statements*, Journal of Business Ethics 2010, 91 (3), 319-327.
[146]Vega, G., *Caveat Emptor: Ethical Chauvinism in the Global Economy*, Journal of Business Ethics, 1997, 16 (12), 1353-1362.
[147]Vitell, S. J. and Hidalgo, E. R., *The Impact of Corporate Ethical Values and Enforcement of Ethical Codes on the Perceived Importance of Ethics in Business: A Comparison of U.S. and Spanish Managers*, Journal of Business Ethics, 2006, 64 (1), 31-43.
[148]Vitell, S. J., Nwachukwu, S. L. and Barnes, J. H., *The Effects of Culture on Ethical Decision-Making: An Application of Hofstede's Typology*, Journal of Business Ethics, 1993, 12 (10), 753-760
[149]Volkema, R. J. and. Fleury, M. T. L., *Alternative Negotiating Conditions and the Choice of Negotiation Tactics: A Cross-cultural Comparison*, Journal of Business Ethics, 2002, 36 (4), 381-398.
[150]Voyer, P. A. and Beamish, P. W., *The Effect of Corruption on Japanese Foreign Direct Investment*, Journal of Business Ethics, 2004, 50 (3), 211-224.
[151]Wan, W. W., Luk, C. L., Yau, O. H., Tse, A. C., Sin, L. Y., Kwong, K. K. and Chow, R. P., *Do Traditional Chinese Cultural Values Nourish a Market for Pirated CDs?*, Journal of Business Ethics, 2009, 88 (1), 185–196.
[152]Wang, C. L. and Lin, X.., *Migration of Chinese Consumption Values: Traditions, Modernization, and Cultural Renaissance*, Journal of Business Ethics, 2009, 88 (3) 399-409.
[153]Weaver, G. R., *Ethics Programs in Global Businesses: Culture's Role in Managing Ethics*, Journal of Business Ethics, 2001, 30 (1), 3-15.
[154]Whitcomb, L. L., Erdener, C. B. and Li, C., *Business Ethical Values in China and the U.S.*, Journal of Business Ethics, 1998, 17 (8), 839-852.
[155]Wimalasiri, J. S., *Contrasts in Moral Reasoning Capacity: The Fijians and the Singaporeans*, Journal of Business Ethics, 2004, 49 (3), 251-272.
[156]Woodbine, G. F., *Moral Choice and the Declining Influence of Traditional Value Orientations within the Financial Sector of a Rapidly Developing Region of the People's Republic of China*, Journal of Business Ethics, 2004, 55 (1), 43-60.
[157]Yang, D., Sonmez, M., Bosworth, D. and Fryxell, G.., *Global Software Piracy: Searching for Further Explacountries*, Journal of Business Ethics, 2009, 87 (2), 269-283.
[158]Zgheib, P. W., *Managerial Ethics: An Empirical Study of Business Students in the American University of Beirut*, Journal of Business Ethics, 2005, 61 (1), 69-78.
[159]Zhang, J., Chiu, R. and Wei, L., *Decision-Making Process of Internal Whistleblowing Behavior in China: Empirical Evidence and Implication*, Journal of Business Ethics, 2009, 88 (1), 25-41.

DOLLARIZATION IN EL SALVADOR REVISITED

Luis Eduardo Rivera-Solis
Dowling College
Oakdale, NY 11769
riveral@dowling.edu

ABSTRACT

The economic crisis in different parts of Latin America has given rise to consideration of dollarization as an option for many of these countries. Those countries that seek to replace their domestic currencies with the US dollar do so with the hope of achieving both growth and economic stability. At the beginning of 2001, El Salvador began a bold experiment of with dollarization. According to many economists, dollarization does nothing to resolve structural and institutional problems which, in many cases, give rise to economic problems. The purpose of this paper is to address the dollarization in El Salvador. This paper will address both the costs and benefits facing the path that El Salvador has chosen

INTRODUCTION

The process where residents of a country utilize the US dollar instead of or alongside the domestic currency is called dollarization. Dollarization may be official or unofficial. It is official if the government of the country accepts the US dollar (or another foreign currency) as the predominant or legal tender currency. It is unofficial whenever the residents of that country hold US dollars, whether it is in the form of foreign currency deposits or paper money in order to protect themselves against high domestic inflation.

Recent currency crises have given impetus to dollarization. Both industrialized (Italy and the United Kingdom 1992) and emerging market countries (Brazil in 1997, Mexico in 1994, and East Asia in 1997) have undergone currency crises Currency crises, especially for emerging markets, can be quite costly. Mexico, for example, experienced a 7 percent decline in real GDP in 1995, while prior to the crises it was experiencing growth rates ranging from 3 to 10 percent... In addition, the crisis in Mexico spilled over into Argentina following the Mexican crisis. In addition contagion this phenomenon is known as contagion.

Under official or full dollarization, the U.S. dollar is the legal tender for all transactions in the economy. Several countries have already officially dollarized. Panama adopted the U.S. dollar as its official currency in 1904, Ecuador dollarized in September 2000 and El Salvador dollarized in January 2001. Unofficial or partial dollarization, which is widespread in Latin America, refers to the process where individuals substitute domestic money with foreign money in order to conduct transactions and protect the purchasing power of their income. [5] Those countries choosing to replace their domestic currencies with the US dollar are seeking to obtain economic growth and stability. By replacing their domestic currencies with the U.S. dollar, countries considering dollarization hope to achieve economic stability and growth, which is the case of El Salvador. [5]

The purpose of this paper is to address the dollarization in El Salvador. This paper will address both the costs and facing the path that El Salvador has chosen. Lastly, it will consider whether or not dollarization does not "turn out to be a "dead end" [3]. In other words, will this policy be sustainable

WHY DOLLARIZE?

In January 2001, El Salvador decided to adopt the US dollar as its currency. However, it took this decision from a position of strength, which is different for the case of Ecuador. In 1999, Ecuador dollarized in order to avoid an economic collapse. According to the finance Minister,

Juan Jose Daboub, this decision was part and parcel of its economic reforms. Given the country's low fiscal deficits and low debt levels, "dollarization reduces interest rates and increases predictability, which is a boost for investment."[2]

Dollarization appears to make sense. El Salvador relies substantially on remittances sent from Salvadorians living in the US. The US is also El Salvador's principal trading partner. It is estimated that there are about two million Salvadorians in the US. It is also estimated that they remitted about $1.97 billion. This represents approximately 13% of El Salvador's GDP. [2]In order to dollarize, there needs to be a government that is seriously committed to maintaining consistency in its economic policy, and important structural and institutional reforms must be put in place. In the case of El Salvador underwent structural and stabilization reforms in which dollarization were part of the process. This was done on the expectation that foreign investment would be promoted.[5] Included in the reforms were trade liberalization, pension reform, and privatization of the banking, telecommunications and electricity distribution industries.

Another important factor is that El Salvador's external-debt –to-GDP ratio is very low (26%) and its total debt ratio of 35% is low. This means that El Salvador needs only a small proportion of its export revenues (10%) to meet its debt obligations. Inflation was 1.4% in 2001, down from as high as 25% in the 1980s. Furthermore, the economy has become more diversified. Coffee accounted for 60% of its exports up until the late 1990's. As a result of maquiladora assembly and export processing plants, that percentage has declined to 8%. These exports have been directed primarily to the US from $790 million in the early 1990's to $2.1 billion in 2001. In addition, El Salvador has attracted foreign investment, which in 2001 rose to $200 million 10% more than in 2000 .[2]

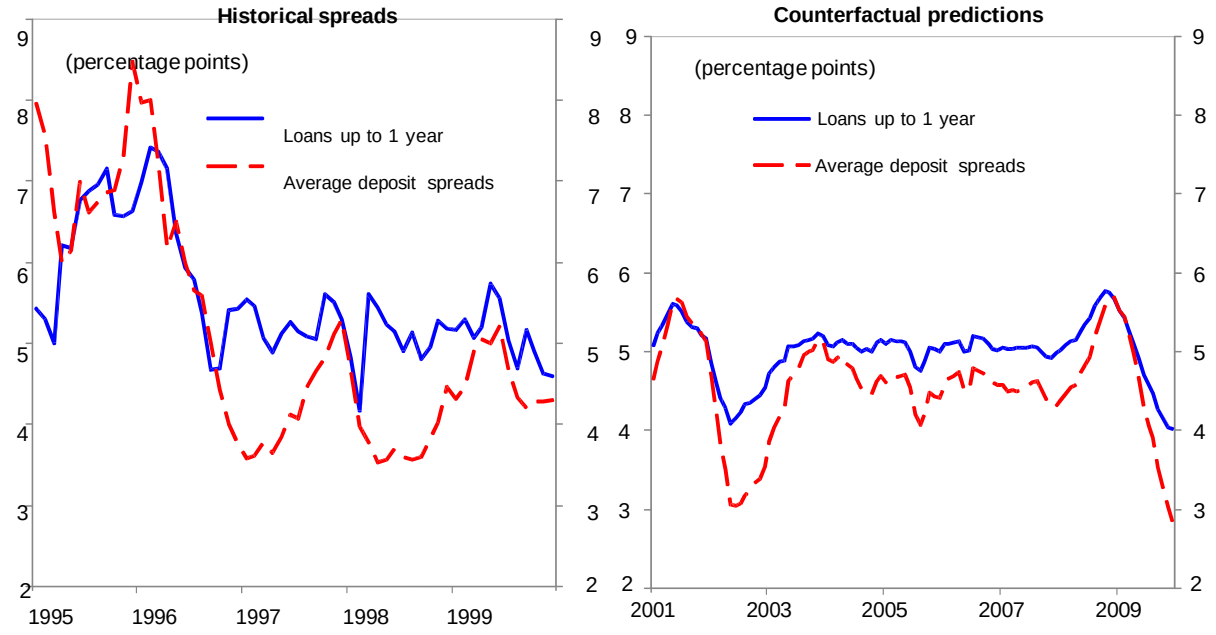
COSTS AND BENEFITS OF DOLLARIZATION

Dollarization has costs and benefits. To begin with, with the money supply will vary with net foreign reserves. This has important economic implications. The central bank cannot serve as "a lender of last resort". Its primary function is relegated to maintaining a sound and efficient financial system. Also, fiscal deficits have to be financed by alternative means as the government is not a position to print money. However, restrictions on monetary and fiscal policy could be seen as a good thing in a region with a past history of hyperinflation. In the event of a recession, if the government tried to institute countercyclical spending, such a move could be interpreted as the loss of the country's commitment to maintaining consistency in its economic policy. Another cost to consider is the loss of seigniorage which is the revenue from issuing currency. "Net seigniorage is the difference between the cost of putting money into circulation and the value of the goods the money will buy which is the revenue that the government receives for creating money."[5] According to Swiston [6, p.11]," the public sector's gains from lower interest rates outweigh the estimated foregone seigniorage revenue by ¼ percent of GDP per year." In addition, seiniorage only transfers resources from the private sector to the public sector. It does not create wealth.

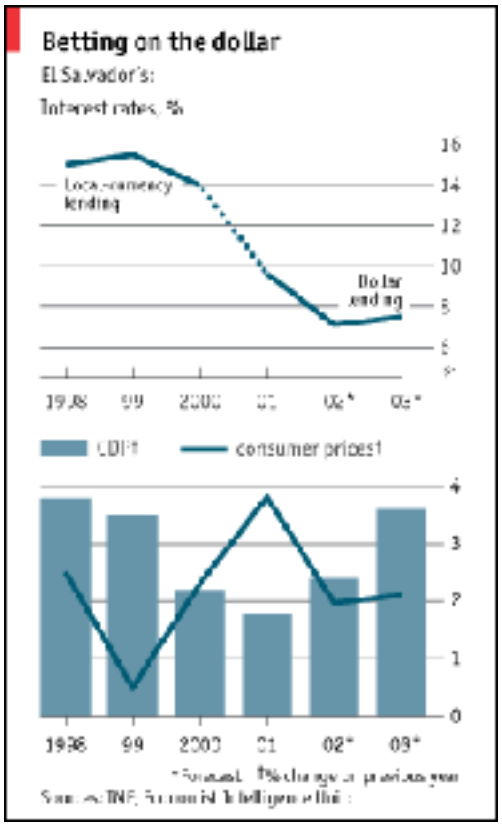
It was expected that interest rates in a fully dollarized economy would decline. This is true for El Salvador as interest rates fell from 14% to 7.5% in 2001. Figure 1 and 2 reveal that since dollarization interest rates have declined. In addition, inflation is relatively low and growth rates have been good. Offsetting the fall in rates could be an increase in country risk or default risk. To date this has not been the case for El Salvador. El Salvador has received investment grade ratings from both Moody's and Fitch ratings. Standard and Poor's has given El Salvador a BB+. It appears that investors find El Salvador as a good investment as it is based on the dollar and they are not subject to currency risk. The present soundness of the economy also

makes El Salvador attractive to investors. [2] The ability of a country to meet its debt obligations affects default risk. To avoid default risk, it is imperative that El Salvador continue to maintain its commitment to debt repayment and fiscal discipline.

Figure .1 *Colón*-Dollar Spreads: History and Counterfactual Predictions



Source: Swiston (2011),



However there have been adverse conditions which have negatively impacted on economic growth, such the increase in oil prices, major earthquakes, and political uncertainty. Because of the high oil prices, shows and estimated real GDP growth or approximately 1.5%. Because of high oil prices, inflation is also expected to be around 5.. An important factor that has helped offset the rise in oil imports is family remittances from abroad. These remittances have aided in the decline of the current account deficit. [4]

Another important benefit of dollarization is the opening up of the financial system. The increased capital mobility resulting from dollarization there is greater competition and efficiency. The financial system becomes more integrated with the rest of the world. The risk here is that the country is exposed to "external shocks, and it can make coping with internal shocks more costly." [5] Should a natural disaster occur, the country would have to depend more heavily on external resources for its reconstruction efforts. Monetary and exchange rate policies would not be available to deal with the negative impact of such external shocks.

It should be pointed out that in the 2004 Financial System Stability Assessment, the IMF Executive Board found that the financial sector in El Salvador was "that the financial sector generally is in good health and that dollarization has strengthened financial stability" [4] It has been argued that the costs of dollarization for El Salvador outweigh the benefits. In their analysis, dollarization increases inequality and hence it has a negative effect on the poor. As the poor have no access loans they really cannot benefit from low interest rates.

CONCLUSION

As pointed out above dollarization has provided both costs and benefits. Some of those benefits include closer integration with international markets and exposure to competition from these markets. The costs of dollarization include the loss of seigniorage and a potential for greater fragility of the banking system. Such fragilities can limit the policy options available to the authorities, as well as put an additional burden on the central bank as lender of last resort." [1] As it was pointed out above, there was a substantial saving for the economy with the reduction of currency risk premiums which are found in interest rates in El Salvador. In addition, Swiston [6] estimated "net interest savings of ½ percent of GDP per year for the Salvadoran private sector and ¼ percent of GDP for the public sector, taking into account the opportunity cost of foregone seigniorage under dollarization. "

The key question that arises is whether or not dollarization is sustainable in the long run. A fully dollarized economy has to rely on foreign currency or international reserves to flow on a continuous basis as well on international market conditions. It has to be able to attract foreign investment and net borrowing in order to maintain economic growth. [5] To be sustainable, dollarization requires fiscal discipline and a sound financial system that can respond to financial crises. In essence, dollarization results in the dismantling of both monetary and fiscal policy.

According to Izurieta [3], "…by dismantling the exchange rate and monetary policy, fiscal policy is no longer a policy tool if the financial stability of the system is to be guarded." If that is the case, policymakers do not have the necessary tools to protect employment and generate income. In the case of El Salvador, dollarization could turn out to be a "dead end".

REFERENCES

[1] Balino, T; Bennett, A ; Borensztein, E. Monetary Policy in Dollarized Economies, *International Monetary Fund Occasional Paper No. 171*, March 15,1999.

[2] Emmot, R., Dollarizing from a Position of Strength*", Latin Finance*, September 2002..

[3] Izurieta, A., "Dollarization: A Dead End?", *Working Paper No. 344*, March 2002, The Levy Economics Institute.

[4] International Monetary Fund, "IMF Executive Board Concludes Article IV Consultation With El Salvador", *Public Information Notice (PIN) No. 05/12*, February 14,2005.

[5] Joint Economic Committee, Basics of Dollarization, *US Congress*, January 2000. Latin American Research Group (LARG), Dollarization in Latin America", *Federal Reserve Bank of Atlanta*

[6] Swiston A., Official Dollarization as Monetary Regime: Its Effects on El Salvador, *IMF Working Paper*, June, 2011).

[7] The Economist, "El Salvador Learns to Love the Greenback", September 26, 2002

A MULTI-CRITERIA PRODUCT RECOVERY MODEL WITH INTELLIGENT PRODUCTS UNDER UNCERTAINTY

Onder Ondemir, Northeastern Univ., Boston, MA, 02115, (617) 373-7635, ondemir.o@neu.edu
Surendra M. Gupta[1], Northeastern Univ., Boston, MA 02115, (617) 373-4846, gupta@neu.edu

ABSTRACT

Advanced remanufacturing-to-order and disassembly-to-order (ARTODTO) systems decrease or almost eliminate uncertainty on the condition of end-of-life products (EOLPs) by utilizing sensors and RFID tags. However, there are several other uncertain aspects of ARTODTO systems that need to be addressed, namely, the system goals. This paper presents a fuzzy multi objective ARTODTO model. The goal of the proposed model is to determine how to process each and every EOLP on hand to meet used product and component demands as well as recycled material demand so that various uncertain system goals are achieved. A numerical case example is considered for illustration purposes.

Keywords: Remanufacturing, Disassembly, Sensors, RFID, Fuzzy Goal Programming

INTRODUCTION

Ever-increasing supply of goods and alarming depletion of virgin resources have led to several environmental problems. These have resulted in several legislations that hold manufacturers responsible for their products even after the point of sale, thus imposing optimal management of end-of-life products (EOLPs) and product recovery.

Product recovery may be of several forms such as remanufacturing, refurbishing, repairing and recycling. All recovery options involve disassembly operations up to a certain level. Disassembly is a labor intensive operation carried out to extract parts from EOLPs for several purposes including elimination of hazardous materials, reusable component recovery, component testing, and content inspection. Of all recovery operations, remanufacturing and disassembly are considered to be the most complex ones. This is mostly due to the lack of information about the quality and quantity of EOLPs and their components. Each EOLP has its own quality condition thus exhibiting unique remanufacturing needs. Hence, finding the EOLPs with minimal recovery costs requires testing the whole EOLP inventory, which can be very expensive. However, emerging information technology devices, such as sensors and radio-frequency identification (RFID) tags, can mitigate EOL recovery decisions.

[1] **Corresponding author**. Laboratory for Responsible Manufacturing (LRM), Department of MIE, Northeastern University, 360 Huntington Avenue, Boston, MA, 02115 USA

Sensors and RFID tags, once incorporated with the products, can monitor the critical components throughout products' economic lives and deliver the collected lifecycle information when the products reach recovery facilities. Complete knowledge on the condition and quantity of EOLPs and remaining life determination eliminate costly preliminary disassembly and inspection operations, and enable optimal remanufacturing planning.

LITERATURE REVIEW

Product recovery

The state of the art on Environmentally Conscious Manufacturing and Product Recovery (ECMPRO) through 1999 is presented by Gungor and Gupta [5]. Ilgin and Gupta [10] complemented this work exploring the state of the art through 2009. One of the most researched topics in ECMPRO is disassembly because of its importance in all recovery operations. Although disassembly problems have been studied under several subtopics, disassembly-to-order (DTO) research exhibits the most relevance to this work. Kongar and Gupta [17] presented a preemptive goal programming (PGP) model of an electronic products disassembly-to-order system considering a variety of physical, financial and environmental constraints and goals. In [20], the same authors proposed a linear physical programming (LPP) model to solve the DTO problem with multiple physical targets. Heuristics, meta-heuristics and expert models have been used in the DTO problem because of the increased complexity especially when multiple criteria and/or periods are considered. Kongar and Gupta [19] presented a multi objective tabu search (TS) algorithm and Gupta *et al.* [6] proposed an artificial neural network model in order to solve the DTO problem.

Sensor and RFID technology

Several publications [3, 11, 12] have appeared on embedded sensors for after-sale monitoring and end-of-life decision making. Yang *et al.* [25] performed field trials to show that sensor-embedded products can enable EOL treatment. Ilgin and Gupta [9] investigated the impact of SEPs on the various performance measures of a multi-product disassembly line using simulation analysis. Via a cost-benefit analysis, the authors showed that SEPs not only reduce the total system cost, but also increase the revenue and profit. Also, in [13], the authors simulated the recovery of sensor-embedded washing machines using a multi-kanban controlled disassembly line. RFID tags have become prominent in environmentally conscious product recovery. Parlikad and McFarlane [22] discussed the positive impacts of RFID on product recovery decisions. Gonnuru [4] proposed an RFID integrated fuzzy based disassembly planning and sequencing model and showed the use of life-cycle information for optimal disassembly decisions. Zhou *et al.* [27] proposed an RFID-based remote monitoring system for enterprise internal production management. Product life cycle data, once captured, allow remaining life time estimation [23, 24]. Herzog *et al*. [8] proved the advantage of using condition-based data in remaining life prediction. Byington *et al.* [2] proposed a data-driven neural network methodology to remaining life predictions for aircraft actuator components.

Fuzzy Optimization in Product Recovery

The notion of fuzzy sets was first introduced by Zadeh [26] in order to accommodate vague (uncertain) information in mathematical operations. Zimmermann [28, 29] investigated the application of fuzzy sets in various areas including optimization and decision making. Researchers investigated the use of fuzzy sets in operation research [28] and developed several

fuzzy goal programming (FGP) approaches for the situations where the decision makers (DM) are uncertain about the goals. A recent book by Jones and Tamiz [16] provide in-depth information on GP and FGP.

Product recovery operations are known to have high degrees of uncertainty. Inderfurth [14] showed the negative effects of return, quality and demand uncertainties on the recovery fraction via a numerical analysis. Hence, deterministic models gave way to fuzzy, stochastic, heuristic, meta-heuristic, and expert models. Barba-Gutierrez and Adenso-Diaz [1] extended ref. [7] using a fuzzy logic approach, incorporating imprecision and subjectivity into the model formulation and solution process. Langella [21] developed a multi period heuristic considering holding costs and external procurement of items. Inderfurth and Langella [15] developed two heuristic procedures to investigate the effect of stochastic yields on the DTO system. Kongar and Gupta [18] presented a fuzzy goal programming model for DTO systems under uncertainty.

ADVANCED REMANUFACTURING-TO-ORDER AND DISASSEMBLY-TO-ORDER SYSTEM

The ARTODTO system is a product recovery system where products are embedded with sensors and RFID tags. The ARTODTO system has a remanufacturing option to meet products demands. Since life-cycle data collected by sensors enable remaining useful life determination, the ARTODTO system can respond to remaining-life based demands. Hence, the problem is to find an optimal set of EOLPs to remanufacture, disassemble, and recycle so that the remaining-life based demands are satisfied while the uncertain goals of the system are achieved as much as possible. The ARTODTO system described is depicted in Figure 1.

Figure 1

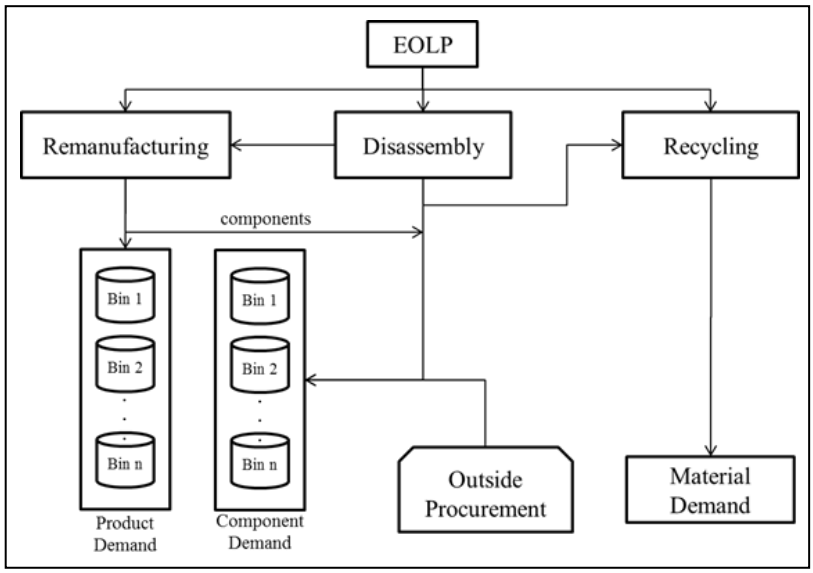


Figure 1: Material and information flows of the ARTODTO system

The proposed ARTODTO model has three fuzzy goals. The first goal is related to the total cost (*TC*) of ARTODTO activities. It is desired to have a total cost value of at most *TC**. However, values above the aspiration level are allowed up to an upper bound with decreasing achievement degrees. The second goal is to attain or exceed the target material sales revenue (*MSR*)* value.

The final objective of the system is related to the customers' satisfaction level (*CSL*) on remaining life of ordered items. Customers' satisfaction level is the difference between the estimated remaining lives of supplied components and the minimum remaining life requirements of orders. The goal is to reach the target customers' satisfaction level (*CSL**) that is defined by the decision makers.

PROPOSED MATHEMATICAL MODEL

The nomenclature used in the mathematical formulation is given in Table 1.

Table 1: Variables and parameters used in the paper

Variables	Definition
CSL	: Customers' satisfaction level;
MSR	: Material sales revenue;
TC	: Total cost;
TDC	: Total disassembly cost;
TRC	: Total recycling cost;
$TRMC$	: Total remanufacturing cost;
def_{ijb}	: 1 if component j in EOLP i is disassembled because of remaining life deficiency and placed in remaining-life-bin b during remanufacturing, zero otherwise;
fr_j	: Number of nonfunctional component js that are recycled;
l_{jb}	: Number of component js purchased for remaining-life-bin;
μ_t	: Achievement level of the t^{th} goal, $t = \{1, 2, 3\}$;
r_{jb}	: Number of component js in remaining-life-bin b that are recycled;
rep_{imjb}	: 1 if a component j from life-bin b needs to be used to remanufacture EOLP i in order to make a product for life-bin m, zero otherwise;
rp_{ij}	: 1 if component j in EOLP i is disassembled during remanufacture, zero otherwise;
sm	: Amount of excess steel sold;
u	: The model objective to be minimized;
w_i	: 1 if EOLP i is recycled, zero otherwise;
$\bar{x}_i$	: 1 if EOLP i is disassembled, zero otherwise;
x_{ijb}	: 1 if component j in EOLP i is disassembled and placed in remaining-life-bin b, zero otherwise;
$\bar{y}_i$	: 1 if EOLP i is remanufactured, zero otherwise;
y_{im}	: 1 if EOLP i is remanufactured to make a product for remaining-life-bin m, zero otherwise;

Parameters	Definition
I	: Set of EOLPs on hand;
B,M	: Sets of remaining-life-bins;
J	: Set of components dealt with;
P_j	: Set of predecessor components of component j;
b, i, j, j', m	: Running numbers;
A	: Destructive disassembly cost factor;
a_{ij}	: 1 if component j of EOLP i is functional;
β_{ij}	: The highest life-bin that component j of EOLP i can be placed in;
c_{jb}	: Outside procurement cost of a component j for life-bin b;
ca_j	: Assembly cost of a component j;
cd_j	: Disassembly cost of a component j;
cin_{ijb}	: 1 when remaining life of component j is adequate for remaining-life-bin b, zero otherwise;
crc_i	: Recycling cost of a component j;
dc_{jb}	: Demand for component js in remaining-life-bin b;
dp_m	: Demand for products in remaining-life-bin m;
dm	: Demand for steel;
dfc_{imj}	: 1 if component j of EOLP i is remaining-life-deficient for life bin m, zero otherwise;
f_{ij}	: 1 if component j of EOLP i is non-functional, zero otherwise;
γ_j	: Steel yield of a component j,
LB_t	: Lower bound of the t^{th} goal;
mis_{ij}	: Binary parameter taking 1 if component j is missing in EOLP i, zero otherwise;
prc	: Unit sales price of steel;
UB_t	: Upper bound of the t^{th} goal;

Goals

Aforementioned goals of the system can be formulated as follows:

$$TC \lesssim TC^* \quad (1)$$

$$MSR \gtrsim MSR^* \quad (2)$$

$$CSL \gtrsim CSL^* \quad (3)$$

Constraints

The first goal is a left-sided fuzzy set, hence the achievement level of the first goal can be written as follows:

$$\mu_1 \leq \frac{UB_1 - TC}{UB_1 - TC^*} \tag{4}$$

Where,

$$TC = TDC + TRMC + TOPC + TRC \tag{5}$$

Each cost term is described below.

Total disassembly cost, *TDC*, is incurred by completely disassembled EOLPs. *TDC* can be formulated as:

$$TDC = \sum_{i\epsilon I, j\epsilon J} \left(\bar{x}_i(a_{ij}cd_j + f_{ij} \propto cd_j) + w_i(a_{ij} + f_{ij})\ \alpha\ cd_j\right) \tag{6}$$

Remanufacturing comprises of the disassembly of broken and remaining-life-time deficient components, and assembly of required ones. Therefore:

$$TRMC = \sum_{i\epsilon I, j\epsilon J} \left[rp_{ij}\left(a_{ij}\left(cd_j + ca_j\right) + f_{ij}\left(\alpha cd_j + ca_j\right) + mis_{ij}ca_j\right)\right] \tag{7}$$

and,

$$y_{im}\left(f_{ij} + mis_{ij} + dfc_{imj}\right) \leq rp_{ij'}, \qquad \forall i, m, \{j, j' | j' \in P_j\} \tag{8}$$

TOPC can be written as follows:

$$TOPC = \sum_{j\in J, b\in B} c_{jb}l_{jb} \tag{9}$$

Total recycling cost, *TRC*, is calculated by summing the costs of product and component recycling. Therefore

$$TRC = \sum_{j\in J} crc_j \left(\sum_{b\in B} r_{jb} + fr_j\right) \tag{10}$$

The achievement level of the second goal can be written as follows:

$$\mu_2 \leq \frac{MSR - LB_2}{MSR^* - LB_2} \tag{11}$$

Material sales revenue (*MSR*) can be calculated by multiplying the amount of recycled material with the unit material sale price. Therefore

$$MSR = prc\left(\sum_{j\in J}\gamma_j\left(\sum_{b\in B} r_{jb} + fr_j\right)\right) \tag{12}$$

Finally, the achievement level of the third goal can be written as follows:

$$\mu_3 \leq \frac{CSL - LB_3}{CSL^* - LB_3} \tag{13}$$

Customers' satisfaction level (*CSL*) is a conceptual notion and calculated as the sum of two terms, namely, the sum of the differences between the highest life bins that components could be placed and the life bins they are actually placed in. Hence,

$$\begin{aligned} CSL = &\sum_{i\in I, j\in J, b\in B} x_{ijb}\left(\beta_{ij} - b\right) + (b - m) \sum_{i\in I, m\in M, j\in J, b\in B} rep_{imjb} \\ &+ \left(\beta_{ij} - m\right) \sum_{i\in I, m\in M, j\in J}\left(a_{ij}y_{im} - \sum_{b\in B} rep_{imjb}\right) \end{aligned} \tag{14}$$

Following equation introduces the preemptive importance of the goals

$$\mu_1 \geq \mu_2 \geq \mu_3 \tag{15}$$

Equation set below ensures that an EOLP in the inventory is disassembled, remanufactured, recycled or left untouched. Thus,

$$\bar{x}_i + \bar{y}_i + w_i \leq 1, \qquad \forall i \tag{16}$$

Each component in an EOLP is extracted and placed in only one life-bin, if the EOLP is disassembled. Therefore

$$\sum_{b\in B} x_{ijb} = \bar{x}_i a_{ij}, \qquad \forall i, j \tag{17}$$

Equation set below ensures that an EOLP is remanufactured to produce only one product and that product is placed in only one product life-bin.

$$\sum_{m\in M} y_{im} = \bar{y}_i, \qquad \forall i \tag{18}$$

Sophisticated product demand is satisfied by remanufactured EOLPs. Thus,

$$\sum_{i\in I} y_{im} = dp_m, \qquad \forall m \tag{19}$$

For each life bin *b* and component *j*, the number of recovered and procured components must be at least equal to the number of demanded components after components used in remanufacturing and recycling operations are taken out. The constraint can be written as follows.

$$\sum_{\{i\in I|cin_{ijb}=1\}}\left(x_{ijb}+def_{ijb}\right)-\sum_{i\in I,m\in M}\left(rep_{imjb}\right)+l_{jb}-r_{jb}\geq dc_{jb},\qquad \forall b,j \tag{20}$$

where,

$$\sum_{\{b\in B|cin_{ijb}=1\}} def_{ijb}=\sum_{m\in M}\left(dfc_{imj}y_{im}\right),\qquad \forall i,j \tag{21}$$

$$\sum_{\{b\in B|cin_{ijb}\neq 1\}} def_{ijb}=0,\qquad \forall i,j \tag{22}$$

Non-functional, missing, and remaining-life-time deficient components must be replaced of using proper components during remanufacturing. Therefore

$$\sum_{\{b\in B,m\in M|b\geq m\}} rep_{imjb}=y_{im}\left(f_{ij}+mis_{ij}+dfc_{imj}\right),\qquad \forall i,j,m \tag{23}$$

$$\sum_{\{b\in B,m\in M|b<m\}} rep_{imjb}=0,\qquad \forall i,j \tag{24}$$

Material demand is satisfied by recycled components. Hence,

$$\sum_{j\in J}\gamma_j\left(\sum_{b\in B}r_{jb}+fr_j\right)-sm=dm, \tag{25}$$

All non-functional components are recycled due to environmental regulations. This is ensured by the following equation:

$$fr_j=\sum_{i\in I}(\bar{x}_i+\bar{y}_i+w_i)f_{ij}+w_i a_{ij},\qquad \forall j \tag{26}$$

Therefore the FGP model can be written as follows:

$$Max.u=\sum_{t=\{1,2,3\}}\mu_t \tag{27}$$

Subject to:

Eq. (4) - (27).

NUMERICAL EXAMPLE

In this section, sensors and RFID tags embedded water heater ARTODTO system is considered to illustrate the proposed model's application. Each heater contains 8 components that are subject to disassembly precedence relationships. Components and products are assumed to fall into three remaining life categories. These categories are called remaining-life bins. The first bin holds those items that have a remaining-life time of less than 4 years, the second bin holds those items whose remaining-life is at least 4 years but less than 6 years. The last bin holds the items having 6 years or more remaining life.

Component disassembly relationships, disassembly, assembly, and recycling costs are given in Table 2, Remaining life based component demands and outside procurement prices are given in Table 3.

Table 2: Component related operational costs

Component	Predecessors	Disassembly Cost ($)	Assembly Cost ($)	Recycling Cost ($)
Outer Door (A)	-	0.50	0.50	0.30
Door+Manifold (B)	-	2.00	2.00	0.26
Burner (C)	A, B	0.15	0.15	-
Thermocouple (D)	C	0.25	0.25	-
Pilot Assembly (E)	D	1.00	1.00	-
Igniter (F)	-	0.15	0.15	-
Gas Valve (G)	A, B, F	1.00	1.00	-
Tank (H)	E, G	0.00	0.00	12.00

Table 3: Component related remaining life specific data

Component	Demand			Procurement Price ($)		
	Bin 1	Bin 2	Bin 3	Bin 1	Bin 2	Bin 3
Outer Door (A)	28	28	13	13.99	13.99	13.99
Door+Manifold (B)	19	18	15	20.99	20.99	20.99
Burner (C)	19	32	18	12.99	27.49	41.99
Thermocouple (D)	6	6	16	7.99	7.99	7.99
Pilot Assembly (E)	31	34	20	9.99	17.99	25.99
Igniter (F)	34	26	13	6.99	11.74	16.49
Gas Valve (G)	54	20	16	74.99	134.49	193.99
Tank (H)	45	33	12	56.24	100.87	145.49

Remaining life based remanufactured product demands are assumed to be 12, 13, and 15 for the bins 1, 2, and 3, respectively.

Water heaters are recycled to obtain steel. Steel yields of outer door, door and tank are 1.00, 1.00 and 50.00 lbs., respectively. Unit sale prices of steel is taken as $0.50/lb. and the demand for steel is 1500 lbs.

Additional data include: $\alpha = 0.80$, $B = \{1,2,3\}$, $I = \{1,2,\dots,200\}$, $J = \{1,2,\dots,8\}$, $M = \{1,2,3\}$. Item level sensor data are populated from a database table. A sample of this data is given in Table 4. "*****" and "0.00" notations indicate missing and non-functional components, respectively.

Aspiration value and upper bound for total cost are given as $900 ($TC^* = 900$) and $1200 ($UB_1 = 1200$). DMs believe that revenue from the material sales should be at least $500 ($LB_2 = 500$), while the aspiration value is $800 ($MSR^* = 500$). Finally, the aspiration value for customers' satisfaction level is set at 750 ($CSL^* = 750$) because DMs aim to meet the component demands with superior components. DMs also believe that minimum acceptable CSL is 500 ($LB_3 = 500$).

Table 4: Remaining-life times (years) of components

	Components							
EOLP #	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**
1	6.67	5.12	5.97	*	6.76	6.43	6.69	7.56
2	6.61	5.51	6.84	5.88	4.86	6.45	*	3.95
3	5.88	4.62	6.78	5.48	6.34	5.84	8.06	5.85
4	7.22	6.47	7.86	5.28	6.22	4.04	4.22	0.00
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .

RESULTS

The model was solved using CPLEX Interactive Optimizer. Optimal achievement values and levels of the goals are given in Table 5. Table 6 shows the serial numbers of the EOLPs that are subject to various operations according to the optimal solution.

Table 5: Goals, tolerances, achievement values and levels

Goals	Lower Bounds	Aspiration values	Upper Bounds	Achievement Value	Achievement level
Total cost	-	$900	$1200	$951.49	0.8284
Material sales revenue	$500	$800	-	$750	0. 8284
Customers' satisfaction level	500	750	-	707	0.8280

It should be noted that the achievement level of the second goal is in fact 0.8333, not 0.8284. The model does not allow this achievement level to be greater than that of the first goal. Therefore, it can be said that any further improvement in this value could hinder the achievement of the first goal.

Table 6: Allocation of EOLPs to recovery operations

Activity	EOLP serial numbers
Disassembled	10, 13, 15, 19, 23, 24, 27, 29, 31, 33, 35, 36, 39, 44, 45, 49, 51, 52, 53, 54, 56, 57, 60, 61, 64, 71, 75, 77, 78, 79, 80, 81, 82, 84, 86, 89, 90, 92, 93, 94, 95, 97, 98, 102, 104, 107, 109, 110, 111, 113, 115, 118, 124, 125, 126, 128, 129, 130, 131, 132, 134, 135, 139, 140, 141, 142, 143, 144, 146, 147, 148, 150, 151, 152, 155, 157, 159, 164, 166, 168, 170, 171, 174, 175, 176, 179, 180, 185, 187, 188, 189, 190, 196, 198, 199, 200 (92 EOLPs)
Remanufactured	1, 3, 5, 7, 9, 12, 17, 21, 22, 25, 28, 30, 37, 41, 46, 47, 55, 59, 62, 67, 68, 73, 76, 87, 91, 103, 106, 117, 119, 123, 127, 133, 136, 138, 167, 172, 173, 184, 195, 197 (40 EOLPs)
Recycled	2, 20, 26, 40, 43, 48, 50, 69, 74, 83, 100, 101, 114, 116, 145, 154, 156, 161, 162, 181, 182, 192 (22 EOLPs)

Solution also provides the detailed optimal remanufacturing plan indicating which EOLP is to be remanufactured and in what bin it is to be placed. A portion of the remanufacturing plan is given in Table 7. According to this table, the model picks EOLP #1 to fulfill the demand for a product

having at least 6 years of remaining life (bin 3). Optimal disassembly plan showing the bin numbers in which disassembled components are placed is also obtained. A sample from the optimal disassembly plan is given in Table 8. For instance, when EOLP #10 is disassembled, outer door and burner are placed in bin 3; pilot assembly, igniter and tank are placed in bin 2; and the other components are placed in bin 1.

Table 7: A portion of remanufacturing plan

EOLP #	Bin
1	3
3	2
5	3
7	2
9	1
12	3
17	3
21	3
22	1
25	3
…	…

Table 8: A portion of the disassembly plan

	Components							
EOLP #	A	B	C	D	E	F	G	H
10	3	1	3	1	2	2	1	2
13	2	2	3	-	2	3	1	2
15	2	1	2	-	2	3	1	1
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
198	1	1	-	1	3	3	1	2
199	1	1	2	3	-	3	1	1
200	1	-	-	3	2	2	1	1

CONCLUSIONS

In this paper, a fuzzy multi-criteria advanced remanufacturing-to-order and disassembly-to-order (ARTODTO) system was developed. The proposed system utilized the life-cycle data that is collected and delivered by sensors and RFID tags, in order to obtain the optimum disassembly, remanufacturing, disposal, recycling and storage plans under uncertain conditions. Optimal solution of the model provided the serial numbers of the products to be subjected to various recovery operations. The goals were related to total system cost (*TC*), material sales revenue (*MSR*), and customers' satisfaction level (*CSL*). The objective function was to maximize the sum of the achievement levels of all goals while accommodating their relative importance. A case example was considered to illustrate the application of the proposed approach and solved using CPLEX Interactive Optimizer. Achievement levels were within the 10% distance from the maximum possible values, thus showing how realistic the goals were set.

REFERENCES

[1] Barba-Gutierrez, Y. and Adenso-Diaz, B., "Reverse Mrp under Uncertain and Imprecise Demand," *International Journal of Advanced Manufacturing Technology*, 40(3-4), pp. 413-424, 2009.

[2] Byington, C. S., Watson, M., and Edwards, D., "Data-Driven Neural Network Methodology to Remaining Life Predictions for Aircraft Actuator Components," in *Proceedings of 2004 IEEE Aerospace Conference*, 2004.

[3] Chang, C.-Y. and Hung, S.-S., "Implementing RFIC and Sensor Technology to Measure Temperature and Humidity inside Concrete Structures," *Construction and Building Materials*, 26(1), pp. 628-637, 2012.

[4] Gonnuru, V. K., "Radio-Frequency Identification (RFID) Integrated Fuzzy Based Disassembly Planning and Sequencing for End-of-Life Products," *Masters Thesis*, Mechanical Engineering, The University of Texas at San Antonio, San Antonio, 2010.

[5] Gungor, A. and Gupta, S. M., "Issues in Environmentally Conscious Manufacturing and Product Recovery: A Survey," *Computers & Industrial Engineering,* 36(4), pp. 811-853, 1999.

[6] Gupta, S. M., Imtanavanich, P., and Nakashima, K., "Using Neural Networks to Solve a Disassembly-to-Order Problem," *International Journal of Biomedical Soft Computing and Human Sciences (Special Issue on Total Operations Management),* 15(1), pp. 67-71, 2009.

[7] Gupta, S. M. and Taleb, K., "Scheduling Disassembly," *International Journal of Production Research,* 32(8), pp. 1857-1866, 1994.

[8] Herzog, M. A., Marwala, T., and Heyns, P. S., "Machine and Component Residual Life Estimation through the Application of Neural Networks," *Reliability Engineering & System Safety,* 94(2), pp. 479-489, 2009.

[9] Ilgin, M. A. and Gupta, S. M., "Comparison of Economic Benefits of Sensor Embedded Products and Conventional Products in a Multi-Product Disassembly Line," *Computers & Industrial Engineering,* 59(4), pp. 748-763, 2010.

[10] Ilgin, M. A. and Gupta, S. M., "Environmentally Conscious Manufacturing and Product Recovery (Ecmpro): A Review of the State of the Art," *Journal of Environmental Management,* 91(3), pp. 563-591, 2010.

[11] Ilgin, M. A. and Gupta, S. M., "Evaluating the Impact of Sensor-Embedded Products on the Performance of an Air Conditioner Disassembly Line," *The International Journal of Advanced Manufacturing Technology,* 53(9-12), pp. 1199-1216, 2010.

[12] Ilgin, M. A. and Gupta, S. M., "Performance Improvement Potential of Sensor Embedded Products in Environmental Supply Chains," *Resources, Conservation and Recycling,* 55(6), pp. 580-592, 2011.

[13] Ilgin, M. A. and Gupta, S. M., "Recovery of Sensor Embedded Washing Machines Using a Multi-Kanban Controlled Disassembly Line," *Robotics and Computer-Integrated Manufacturing,* 27(2), pp. 318-334, 2011.

[14] Inderfurth, K., "Impact of Uncertainties on Recovery Behavior in a Remanufacturing Environment: A Numerical Analysis," *International Journal of Physical Distribution & Logistics Management,* 35(5), pp. 318-336, 2005.

[15] Inderfurth, K. and Langella, I., "Heuristics for Solving Disassemble-to-Order Problems with Stochastic Yields," *OR Spectrum,* 28(1), pp. 73-99, 2006.

[16] Jones, D. and Tamiz, M., *Practical Goal Programming*. New York: Springer, 2010.

[17] Kongar, E. and Gupta, S. M., "A Multi-Criteria Decision Making Approach for Disassembly-to-Order Systems," *Journal of Electronics Manufacturing,* 11(2), pp. 171-83, 2002.

[18] Kongar, E. and Gupta, S. M., "Disassembly to Order System under Uncertainty," *Omega,* 34(6), pp. 550-561, 2006.

[19] Kongar, E. and Gupta, S. M., "A Multiple Objective Tabu Search Approach for End-of-Life Product Disassembly," *International Journal of Advanced Operations Management,* 1(2-3), pp. 177-202, 2009.

[20] Kongar, E. and Gupta, S. M., "Solving the Disassembly-to-Order Problem Using Linear Physical Programming," *International Journal of Mathematics in Operational Research,* 1(4), pp. 504-531, 2009.
[21] Langella, I. M., "Heuristics for Demand-Driven Disassembly Planning," *Computers & Operations Research,* 34(2), pp. 552-577, 2007.
[22] Parlikad, A. K. and McFarlane, D., "RFID-Based Product Information in End-of-Life Decision Making," *Control Engineering Practice,* 15(11), pp. 1348-1363, 2007.
[23] Rugrungruang, F., "An Integrated Methodology for Assessing Physical & Technological Life of Products for Reuse," *PhD Thesis*, Life Cycle Engineering & Management Research Group, School of Mechanical and Manufacturing Engineering, The University of New South Wales, New South Wales, Australia, 2008.
[24] Wang, W. and Zhang, W., "An Asset Residual Life Prediction Model Based on Expert Judgments," *European Journal of Operational Research,* 188(2), pp. 496-505, 2008.
[25] Yang, X., Moore, P., and Chong, S. K., "Intelligent Products: From Lifecycle Data Acquisition to Enabling Product-Related Services," *Computers in Industry*, 60(3), pp. 184-194, 2009.
[26] Zadeh, L. A., "Fuzzy Sets," *Information and Control,* 8(3), pp. 338-353, 1965.
[27] Zhou, S., Ling, W., and Peng, Z., "An RFID-Based Remote Monitoring System for Enterprise Internal Production Management," *International Journal of Advanced Manufacturing Technology,* 33(7-8), pp. 837-844, 2007.
[28] Zimmermann, H. J., "Using Fuzy Sets in Operational Research," *European Journal of Operational Research,* 13(3), pp. 201-216, 1983.
[29] Zimmermann, H. J., *Fuzzy Set Theory--and Its Applications*. Norwell, MA: Kluwer Academic Publishers, 2001.

A PARTICLE SWARM OPTIMIZATION ALGORITHM FOR SOLVING DISASSEMBLY LINE BALANCING PROBLEM

Can B. KALAYCI, Northeastern University
Department of Mechanical and Industrial Engineering, 334 Snell Engineering Center
360 Huntington Avenue, Boston, Massachusetts 02115 U.S.A.
canberkkalayci@gmail.com, (617)-513-2951
Surendra M. Gupta, Northeastern University
Department of Mechanical and Industrial Engineering, 334 Snell Engineering Center
360 Huntington Avenue, Boston, Massachusetts 02115 U.S.A.
gupta@neu.edu, (617)-373-4846

ABSTRACT

In this paper, we consider a disassembly line balancing problem (DLBP) that concerns with the assignment of disassembly tasks to a set of ordered disassembly workstations while satisfying the disassembly precedence constraints and optimizing the effectiveness of several measures. Since the complexity of DLBP increases with the number of parts in the product, it is significant that an efficient methodology be developed. A new approach based on the particle swarm optimization (PSO) algorithm is proposed to solve the DLBP. A case example is considered to demonstrate the functionality of the algorithm.

Keywords: Disassembly Line Balancing, Particle Swarm Optimization, Metaheuristics

1. INTRODUCTION

Product recovery seeks to obtain materials and parts from old or outdated products through recycling and remanufacturing in order to minimize the amount of waste sent to landfills. Gungor and Gupta [7], Ilgin and Gupta [11] provide an extensive review of product recovery. The first crucial and the most time consuming step of product recovery is disassembly. Disassembly is defined as the systematic extraction of valuable parts and materials from discarded products through a series of operations to use at remanufacturing or recycling after appropriate cleaning and testing operations. Disassembly operations can be performed at a single workstation, in a disassembly cell or on a disassembly line. Although a single workstation and disassembly cell are more flexible, the highest productivity rate is provided by a disassembly line and hence is the best choice for automated disassembly processes, a feature that will be essential in the future disassembly systems [8]. Disassembly operations have unique characteristics and cannot be considered as the reverse of assembly operations. The quality and quantity of components used in the stations of an assembly line can be controlled by imposing strict conditions. However, there are no such conditions of EOL products moving on a disassembly line. In a disassembly environment, the flow process is divergent; a single product is broken down into many subassemblies and parts while the flow process is convergent in an assembly environment. There is also a high degree of uncertainty in the structure, quality, reliability and

the condition of the returned products in disassembly. Additionally, some parts of the product may be hazardous and may require special handling that will affect the utilization of disassembly workstations. Since disassembly tends to be expensive, disassembly line balancing becomes significant in minimizing resources (time and money) invested in disassembly and maximizing the level of automation.

Disassembly Line Balancing Problem (DLBP) is a multi-objective problem that is described in [9] and has mathematically been proven to be NP-complete in [20] making the goal to achieve the optimal balance computationally expensive. Exhaustive search works well enough in obtaining optimal solutions for small sized instances; however its exponential time complexity limits its application on the large sized instances. An efficient search method needs to be employed to attain a (near) optimal condition with respect to objective functions. Although some researchers have formulated the DLBP using mathematical programming techniques [2] [3] [4] [6] [17], it quickly becomes unsolvable for a practical sized problem due to its combinatorial nature. For this reason, there is an increasing need to use metaheuristic techniques such as genetic algorithms (GA) [12] [20], ant colony optimization (ACO) [1] [5] [18] [19] [21] [24], simulated annealing (SA) [15], tabu search (TS) [13] and artificial bee colony (ABC) [14]. See [22] for more information on DLBP. In this paper, particle swarm optimization (PSO) approach is proposed to solve the multi-objective DLBP.

The rest of the paper is organized as follows: In Section 2, problem definition and formulation is given. Section 3 describes the proposed PSO algorithm for the multi-objective DLBP. The computational experience to evaluate its performance on a numerical example is provided in Section 4. Finally some conclusions are pointed out in Section 5.

2. PROBLEM DEFINITION AND FORMULATION

Problem assumptions include the following: A single product type is to be disassembled on a disassembly line; the supply of the end-of-life product is infinite; the exact quantity of each part available in the product is known and constant; a disassembly task cannot be divided between two workstations; each part has an assumed associated resale value which includes its market value and recycled material value; disassembly tasks are to be assigned to a sequence of workstations without violating precedence relationships among the tasks; and complete disassembly is performed on the product. Notation used in the mathematical formulation is given as follows:

CT	Cycle time; maximum time available at each workstation
d_k	Demand; quantity of part k requested
h_k	Binary value; 1 if part k is hazardous, else 0
IP	Set (k_1, k_2) of parts such that part k_1 must precede part k_2
j	Workstation count (1,…, NWS)
k	Part identification (1,…, n)
n	The number of parts for removal
N	The set of natural numbers
NWS	Number of workstations required for a given solution sequence

NWS^*	Minimum possible number of workstations for *PRT*
PRT_k	Part removal time required for k^{th} part
PS_k	k^{th} part in a solution sequence
ST_j	Station time; total processing time requirement in workstation *j*
x_{kj}	Task assignments to work stations; 1 if part *k* is assigned to workstation *j*, else 0

Based on concept and assumptions made in [20], the mathematical formulation of our DLBP is given as follows:

$$\min f_1 = NWS \tag{1}$$

$$\min f_2 = \sum_{j=1}^{NWS} (CT - ST_j)^2 \tag{2}$$

$$\min f_3 = \sum_{k=1}^{n} (k \cdot h_{PS_k}), \quad h_{PS_k} = \begin{cases} 1, \text{ hazardous} \\ 0, \text{ otherwise} \end{cases} \tag{3}$$

$$\min f_4 = \sum_{k=1}^{n} (k \cdot d_{PS_k}), \quad d_{PS_k} \in N, \forall PS_k \tag{4}$$

Subject to:

$$\left\lceil \frac{\sum_{k=1}^{n} PRT_k}{CT} \right\rceil \leq NWS^* \leq n \tag{5}$$

$$\sum_{j=1}^{NWS} ST_j \leq CT, \quad j = 1, 2, ..., NWS \tag{6}$$

$$\sum_{j=1}^{NWS} x_{kj} = 1, \quad k = 1, ..., n \tag{7}$$

$$x_{aj} \leq \sum_{j=1}^{NWS} x_{kj}, \quad \forall (a,k) \in IP \tag{8}$$

In the multi-objective optimization formulation given above, equation (1) aims to minimize the number of workstations in the disassembly line. Equation (2) aims to minimize the total idle time across all workstations and aims to make the idle times at all workstations similar to each other (where $f_2 = 0$ represents perfect balance), equation (3) aims to remove hazardous components as early as possible and equation (4) aims to remove high demand components before low demand components in the disassembly process. Constraint (5) guarantees that the

number of work stations with a workload does not exceed the permitted number. Constraint (6) ensures that the work content of a workstation cannot exceed the cycle time. Constraint (7) ensures that all tasks are assigned to at least and at most one workstation (the complete assignment of each task). Equation (8) imposes the restriction that all the disassembly precedence relationships between tasks should be satisfied.

3. PROPOSED PARTICLE SWARM OPTIMIZATION APPROACH

Since DLBP falls into the NP-Complete class of combinatorial optimization problems, when the problem size increases, the solution space is exponentially increased and an optimal solution in polynomial time cannot be found as it can be time consuming for optimum seeking methods to obtain an optimal solution within this vast search space. Therefore, it is necessary to use alternative methods in order to reach (near) optimal solutions faster. For this reason, a fast and effective PSO approach is used to solve the problem.

PSO algorithm is a swarm intelligence based metaheuristic approach proposed by Kennedy and Eberhart [16]. Its biological inspiration is based on the metaphor of social interaction and communication in composed organisms, such as flock of birds or school of fishes. Scientists found that the synchronized behavior of composed organisms was through maintaining optimal distances between individual members and their neighbors. In order to search for food, members adjust their positions according not only to their own experience, but also to the experience of other particles. A member may modify its position with a velocity towards the best position so far achieved by the member itself and the best solution so far achieved by any of the other exploring members. Information is socially shared among individuals to direct the search towards the best position in the solution space. Thus, velocity plays the important role of adjusting each other for the optimal distance. This idea is the main principle of PSO.

In the PSO method, each member is called a particle. Each particle has a position and moves around in the multi-dimensional space with a velocity. At any instant of PSO, a particle changes its velocity and position by exploiting the best position attained so far by itself and by any other particle of the swarm. The best experiences of the particle itself and its neighbors are memorized for updating the velocity and position. In general, it can be said that the behavior of every particle is a compromise between its individual memory and a collective memory. Similar to GA, PSO is also an optimization algorithm based on population. The PSO algorithm starts from an initial swarm of potential solutions, and all members of entire population are maintained through the search procedure. Each particle, flies through the solution space of the optimization problem searching for the (near) optimum solution and thus its position represents a potential solution for the problem. PSO has some appealing features, including easy implementation, few parameters to tune, and fast convergence rate.

3.1 Solution Representation

One of the most important decisions in designing a metaheuristic lies in deciding how to represent solutions and relate them in an efficient way to the searching space. Also, solution representation should be easy to decode to reduce the cost of the algorithm. In the proposed PSO

algorithm, permutation based representation is used, so elements of a solution string are integers. Each element represents a task assignment to work station. The value of the first element of the array shows which task is assigned to workstations first, the second value shows which task is assigned second and so on. For example, if there are 8 tasks to be assigned to workstations then the length of the solution string is 8. Figure 3.1 illustrates assignment of tasks to workstations as an example.

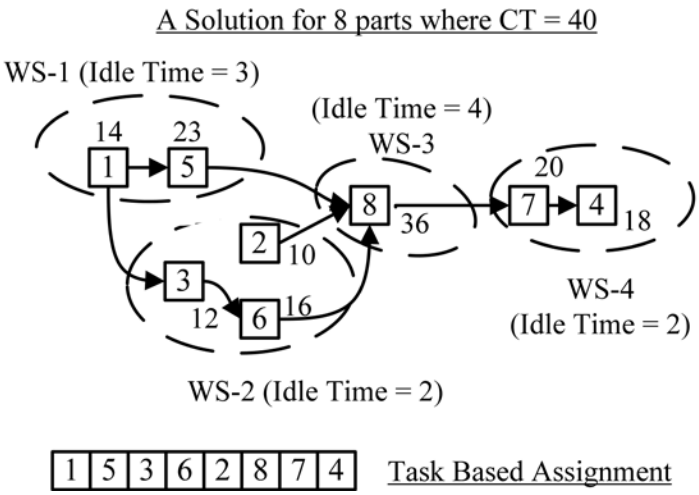


Figure 3.1 Assignment of tasks to workstations

3.2 Modified Continuous Representation

Because of the characteristics of PSO, there seem to be difficulties in deploying permutation type of representation. Real values need to be redefined to make them applicable on a permutation of tasks [23].

Consider the sample task to position representation illustrated in Table 3.1. To construct the continuous version of this representation, we first need to generate 8 (as many as the number of tasks to be assigned) random number between (0, 1), then these numbers will be sorted. Suppose the numbers shown in the Table 3.2 are the random numbers generated. Thus, to build the continuous representation, the random numbers are placed to location sequence according to the sequence of tasks to be assigned. As it can be seen in Table 3.3, the smallest random number is assigned to task 1; the second smallest random number is assigned to task 5 and so on. Finally the continuous representation will be obtained (Table 3.3). The reverse approach is performed to decode the real-coded representation into a solution. In other words, the smallest values in the real-coded representation will be assigned to first location sequence; the next smallest value will be assigned to second location sequence and so on.

Table 3.1 Task to position representation for DLBP

1	2	3	4	5	6	7	8	Location in sequence
1	5	3	6	2	8	7	4	Tasks to be assigned

Table 3.2 A set of random numbers

1	2	3	4	5	6	7	8	Random numbers' sequence
.45	.37	.15	.22	.05	.67	.91	.58	Random numbers
.05	.15	.22	.37	.45	.58	.67	.91	Sorted random numbers

Table 3.3 Continuous representation of Table 1

1	2	3	4	5	6	7	8	Location in sequence
.05	.45	.22	.91	.15	.37	.67	.58	Continuous representation

3.3 Initialization

Each particle presents a potential solution to the optimization problem flying in the *n* (number of parts) dimensional search space, modifying (iteration by iteration) its position according to its own best position in history and that of its neighbor particles.

Each particle *i* has three attributes: its current position $x_{i,g}$, its personal best position achieved so far $x_{i,g}^{pbest}$ and its current velocity $v_{i,g}$. Each one of these attributes is an *n* dimensional parameter vector. The index *g* denotes the iteration number of the algorithm. The initial population (*g*=0), $POP_0 = \left\{ x_{1,0}, x_{2,0}, ..., x_{ps,0} \right\}$ is taken uniformly distributed in the search region using the following formula given in equation (9).

$$x_{i,0}^{k} = x_{\min} + r \times \left(x_{\max} - x_{\min} \right) \tag{9}$$

Where $x_{i,g}^{k}$ is the position value of the *i*th $(i = 1, 2, ..., ps)$ particle with respect to *k*th $(k = 1, 2, ..., n)$ dimension, $x_{\min}$ and $x_{\max}$ are user defined bounds and *r* is a uniform number in (0,1). Similarly the initial velocities of the particles are generated using the formula given in equation (10).

$$v_{i,0}^{k} = v_{\min} + r \times (v_{\max} - v_{\min}) \tag{10}$$

With $v_{\min}$ and $v_{\max}$ being the user defined fixed bounds and *r* a uniform random number in (0,1).

After randomly generating initial positions of the particles, the continuous representation is converted to permutation based feasible line balancing solution as it is described in Section 3.2.

The strategy of building a feasible balancing solution is the key issue to solve the DLBP. We use station-oriented procedure for a solution constructing strategy in which solutions are generated by filling workstations successively one after the other [5]. The procedure is initiated by the opening of a first station. Then, tasks are successively assigned to this station until more tasks cannot be assigned and a new station is opened. In each iteration, a task is randomly chosen from the set of candidate tasks to assign to the current station. When no more tasks may be assigned to the open station, this is closed and the following station is opened. The procedure finalizes when there are no more tasks left to assign.

In order to describe the process to build a feasible balancing solution, available task and candidate task are defined as follows: A task is an available task if and only if it has not already been assigned to a workstation and all of its predecessors have already been assigned to a workstation. A task is a candidate task if and only if it belongs to the set of available task and the idle time of current workstation is higher than or equal to the processing time of the task.

The generation procedure of a feasible balancing solution is given as follows:

Step 1: According to the precedence constraints construct the available task set.

Step 2: According to the cycle time construct the candidate task set.

Step 3: If the set of candidate task is null, go to step 5.

Step 4: Select the task with the highest assignment priority from the candidate task set and assign the task to the current workstation; go back to step 1.

Step 5: If the set of available task is null, go to step 7.

Step 6: Open a new workstation, go back to step 1.

Step 7: Stop the procedure.

3.4 PSO iteration

At each iteration *g*, all particles in *POP* are targeted for replacement. This is achieved by performing the following steps:

1: For each particle $i (i = 1, 2, ..., ps)$

1.1: Evaluate its objective functions

$$F(x_{i,g}) = \left[f_1(x_{i,g}), f_2(x_{i,g}), f_3(x_{i,g}), f_4(x_{i,g}) \right]$$

1.2: Determine its personal best position $x_{i,g}^{pbest}$ as in the following:

$$\text{if } F(x_{i,g}) < F(x_{i,g}^{pbest}) \Rightarrow F(x_{i,g}^{pbest}) = F(x_{i,g}) = \left[f_1(x_{i,g}), f_2(x_{i,g}), f_3(x_{i,g}), f_4(x_{i,g}) \right]$$

2: Determine the global best position x_g^{gbest} corresponding to the best objective function value among the population of the particles (i.e., the whole swarm).

3: For each particle i $(i = 1, 2, ..., ps)$ update its velocity $v_{i,g}$ as in the following:

$$v_{i,g} = c_1 r_1 \left(x_{i,g}^{pbest} - x_{i,g} \right) + c_2 r_2 \left(x_{i,g}^{gbest} - x_{i,g} \right) + Iw_g v_{i,g} \quad (11)$$

Where c_1 and c_2 are cognitive and social parameters respectively, r_1 and r_2 are uniform random numbers drawn in (0,1), c_1 is a weight factor representing the attraction toward $x_{i,g}^{pbest}$, while c_2 is the attraction toward $x_{i,g}^{gbest}$. Both of them are defined to be constants during the execution of the algorithm. Iw_g in Equation (11) is the inertia weight factor which gadgets the effect of the old velocity on to the new one. Iw_g is updated by

$$Iw_g = \Theta \times Iw_{g-1} \quad (12)$$

4: For each particle i $(i = 1, 2, ..., ps)$ calculate its new position using the relation:

$$x_{i,g} = x_{i,g-1} + v_{i,g} \quad \text{for } g > 0 \quad (13)$$

5: Apply local search using SWAP or INSERT operators
6: Repeat steps (1)–(5) until *g* exceeds a maximum (user defined) generations.

3.5 Local Search

As to the permutation based neighborhood structure, INSERT (inserting a task to a different work station) and SWAP (interchanging two tasks) operators are used to produce neighboring solutions such that the new neighboring solutions are ensured to be feasible. By guaranteeing feasibility in each operation, the necessity of the repair function is prevented. The INSERT operator of a permutation is defined by removing a job from its original position and inserting it into another position whereas the SWAP operator produces a neighbor by interchanging two randomly selected tasks in the different positions at workstations while satisfying the precedence constraints. To enrich the neighborhood structure and diversify the population, four neighboring approaches, based on the INSERT or SWAP operator, are separately utilized. Examples for SWAP and INSERT operators are given in Figure 3.2 and Figure 3.3 respectively.

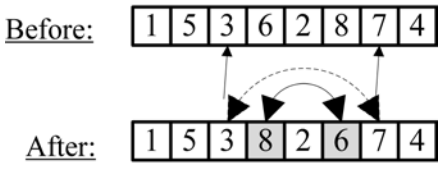


Figure 3.2 SWAP Operation

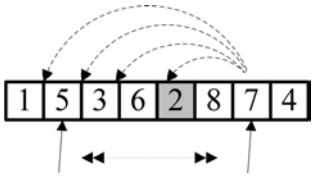


Figure 3.3 INSERT Operation

4. Numerical Results

In this section a numerical example from literature [10] is used to illustrate the effectiveness of the proposed PSO algorithm. Here the objective is to completely disassemble a 25 part cellular phone while satisfying several precedence relationships. Disassembly line operates at a speed which allows *CT*=18 seconds for each workstation to perform its required disassembly tasks. This example consists of the data for the disassembly product as shown in Table 4.1 and Figure 4.1. Parameter values used in the PSO algorithm are as the following: *ps* = 250, $c_1 = 2$, $c_2 = 1.5$, $Iw = .9$, $\Theta = .95$, $x_{\min} = 0$, $x_{\max} = 1$, $v_{\min} = 0$, $v_{\max} = 1$, *generation* = 50.The proposed algorithm was coded in MATLAB and tested on Intel Core2 1.79 GHz processor with 3GB RAM. After 30 PSO runs, all of the PSO solutions found the optimal number of workstations $f_1 = 9$ [10]. From the literature, the minimal value for the measure of balance $f_2 = 9$ [5] is also achieved by the proposed PSO. The hazardous part measure (f_3) averaged 83.5 (ranging from a best result of 80 and a worst result of 90). The high demand part removal measure (f_4) averaged 889.07 (with a best result of 857 and a worst result of 965). A typical PSO solution using the cellular telephone instance is given in Figure 4.2. Computational experience with the algorithm on the benchmark data set has shown that the algorithm performs remarkably well.

Table 4.1 Knowledge base of cellular telephone instance

Task	Time	Hazardous	Demand
1	3	Yes	4
2	2	Yes	7
3	3	No	1
4	10	No	1
5	10	No	1
6	15	No	1
7	15	No	1
8	15	No	1
9	15	No	1
10	2	No	2
11	2	No	1
12	2	Yes	4
13	2	No	1
14	2	No	1
15	2	No	1
16	2	No	1
17	2	No	2
18	3	No	2
19	18	Yes	8
20	5	No	1
21	1	No	4
22	5	No	6
23	15	Yes	7
24	2	No	1
25	2	Yes	4

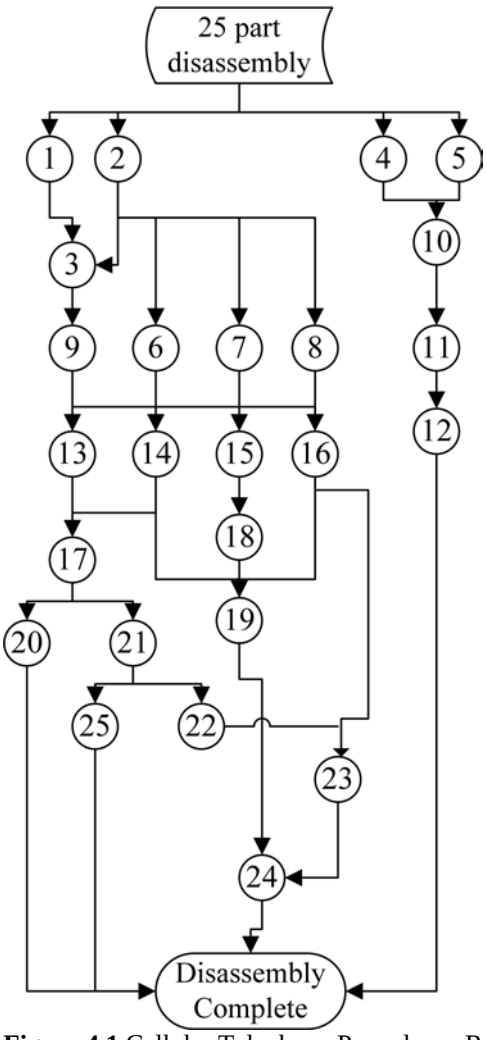


Figure 4.1 Cellular Telephone Precedence Relationships

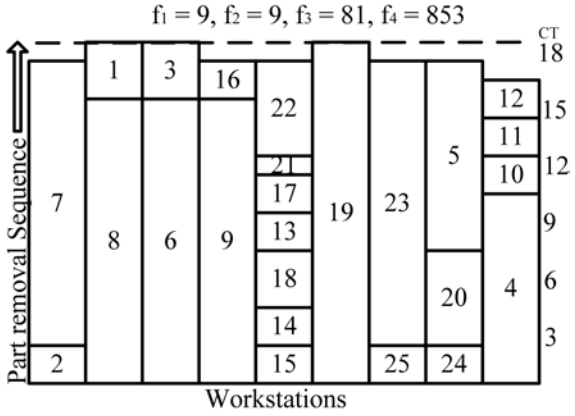


Figure 4.2 A typical solution using the cellular telephone instance

5. Conclusions

The main objective of this paper was to introduce a PSO approach to solve the DLBP with multiple objectives. The approach combined the PSO with a local search procedure. To the best of knowledge of the authors, there is no published study that uses PSO to solve DLBP. Here, PSO was proposed to solve the problem which aimed to minimize the number of disassembly workstations, minimize the total idle time of all workstations by ensuring similar idle time at each workstation, maximize the removal of hazardous components as early as possible in the disassembly sequence and maximize the removal of high demand components before low demand components. A case study from the literature on the disassembly of a 25 part mobile example was used to ascertain the performance of the proposed approach. The PSO approach was able to find (near) optimal solution(s) very quickly. Computational experience with the algorithm of disassembly line balancing problem showed that algorithm worked remarkably well to obtain (near) optimal solutions.

REFERENCES

[1]Agrawal, S., & Tiwari, M. K. (2008). A collaborative ant colony algorithm to stochastic mixed-model U-shaped disassembly line balancing and sequencing problem. *International Journal of Production Research, 46*, 1405 - 1429.

[2]Altekin, F. T., & Akkan, C. (2011). Task-failure-driven rebalancing of disassembly lines. *International Journal of Production Research*, 1-22.

[3]Altekin, F. T., Kandiller, L., & Ozdemirel, N. E. (2004). Disassembly line balancing with limited supply and subassembly availability. In S. M. Gupta (Ed.), *Environmentally Conscious Manufacturing Iii* (Vol. 5262, pp. 59-70). Bellingham: Spie-Int Soc Optical Engineering.

[4]Altekin, F. T., Kandiller, L., & Ozdemirel, N. E. (2008). Profit-oriented disassembly-line balancing. *International Journal of Production Research, 46*, 2675-2693.

[5]Ding, L.-P., Feng, Y.-X., Tan, J.-R., & Gao, Y.-C. (2010). A new multi-objective ant colony algorithm for solving the disassembly line balancing problem. *The International Journal of Advanced Manufacturing Technology, 48*, 761-771.

[6]Goksoy, E. (2010). *Disassembly Line Balancing Problem with Fixed Number of Workstations and Finite Supply*. Unpublished The Degree of Master of Science, Middle East Technical University, Ankara.

[7]Gungor, A., & Gupta, S. M. (1999). Issues in environmentally conscious manufacturing and product recovery: a survey. *Computers & Industrial Engineering, 36*, 811-853.

[8]Gungor, A., & Gupta, S. M. (2001). A solution approach to the disassembly line balancing problem in the presence of task failures. *International Journal of Production Research, 39*, 1427-1467.

[9]Güngör, A., & Gupta, S. M. (2002). Disassembly line in product recovery. *International Journal of Production Research, 40*, 2569 - 2589.

[10]Gupta, S. M., Erbis, E., & McGovern, S. M. (2004). Disassembly sequencing problem: A case study of a cell phone. In S. M. Gupta (Ed.), *Environmentally Conscious Manufacturing IV* (Vol. 5583, pp. 43-52). Bellingham: SPIE-International Society for Optical Engineering.

[11]Ilgin, M. A., & Gupta, S. M. (2010). Environmentally conscious manufacturing and product recovery (ECMPRO): A review of the state of the art. *Journal of Environmental Management, 91*, 563-591.

[12]Kalayci, C. B., & Gupta, S. M. (2011). A hybrid genetic algorithm approach for disassembly line balancing. In *Proceedings of the 42nd Annual Meeting of Decision Science Institute (DSI 2011)*. Boston, MA, USA.

[13]Kalayci, C. B., & Gupta, S. M. (2011). Tabu search for disassembly line balancing with multiple objectives. In *41st International Conference on Computers and Industrial Engineering (CIE41)*. University of Southern California, Los Angeles, USA.

[14]Kalayci, C. B., Gupta, S. M., & Nakashima, K. (2011). Bees Colony Intelligence in Solving Disassembly Line Balancing Problem. In *Proceedings of the 2011 Asian Conference of Management Science and Applications (ACMSA2011)* (pp. 34-41). Sanya, Hainan, China.

[15]Kalayci, C. B., Gupta, S. M., & Nakashima, K. (2011). A Simulated Annealing Algorithm for Balancing a Disassembly Line. In *Proceedings of the Seventh International Symposium on Environmentally Conscious Design and Inverse Manufacturing (EcoDesign 2011)* (pp. 713-718). Kyoto, Japan.

[16]Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. In *Neural Networks, 1995. Proceedings., IEEE International Conference on* (Vol. 4, pp. 1942-1948 vol.1944).

[17]Koc, A., Sabuncuoglu, I., & Erel, E. (2009). Two exact formulations for disassembly line balancing problems with task precedence diagram construction using an AND/OR graph. *IIE Transactions, 41*, 866-881.

[18]McGovern, S. M., & Gupta, S. M. (2005). Uninformed and probabilistic distributed agent combinatorial searches for the unary NP-complete disassembly line balancing problem. In S. M. Gupta (Ed.), *Environmentally Conscious Manufacturing V* (1 ed., Vol. 5997, pp. 81-92). Boston, MA, USA: SPIE.

[19]McGovern, S. M., & Gupta, S. M. (2006). Ant colony optimization for disassembly sequencing with multiple objectives. *The International Journal of Advanced Manufacturing Technology, 30*, 481-496.

[20]McGovern, S. M., & Gupta, S. M. (2007). A balancing method and genetic algorithm for disassembly line balancing. *European Journal of Operational Research, 179*, 692-708.

[21]McGovern, S. M., & Gupta, S. M. (2007). Combinatorial optimization analysis of the unary NP-complete disassembly line balancing problem. *International Journal of Production Research, 45*, 4485-4511.

[22]McGovern, S. M., & Gupta, S. M. (2011). *The Disassembly Line: Balancing and Modeling*. New York: McGraw Hill.

[23]Rahimi-Vahed, A. R., Mirghorbani, S. M., & Rabbani, M. (2007). A hybrid multi-objective particle swarm algorithm for a mixed-model assembly line sequencing problem. *Engineering Optimization, 39*, 877 - 898.

[24]Tripathi, M., Agrawal, S., Pandey, M. K., Shankar, R., & Tiwari, M. K. (2009). Real world disassembly modeling and sequencing problem: Optimization by Algorithm of Self-Guided Ants (ASGA). *Robotics and Computer-Integrated Manufacturing, 25*, 483-496.

SETUP COST REDUCTION IN AN EOQ MODEL WITH NONLINEAR HOLDING COST

Farrokh Nasri, Frank G. Zarb School of Business, Hofstra University, Hempstead, NY 11590, 516-463-4244, Farrokh.Nasri@hofstra.edu

Javad Paknejad, Frank G. Zarb School of Business, Hofstra University, Hempstead, NY 11590, 516-463-4244, M.J.Paknejad@hofstra.edu

John F. Affisco, Frank G. Zarb School of Business, Hofstra University, Hempstead, NY 11590, 516-463-4244, John.F.Affisco@hofstra.edu

ABSTRACT

Many variants of the traditional EOQ model have been developed as a result of relaxing some of its assumptions. These variants include a class of EOQ models that treat holding cost as a function of the amount of time that an item is held in inventory. A second class of EOQ models treat the setup cost as a function of capital expense. In this paper we utilize the basic framework of an EOQ model with nonlinear holding cost that has appeared in the literature and combine it with the strategy of improving flexibility, through setup reduction. Analytical and numerical results are presented.

Keywords: Inventory; flexibility improvement.

1. INTRODUCTION

The classical EOQ model is based on a series of highly restrictive assumptions among which are constant deterministic demand, constant holding costs across time, instantaneous replenishment, and constant order costs per order or replenishment cycle. These assumptions limit the applicability of the model to the actual inventory systems that one encounters in organizations. In recognition of this fact, researchers in recent years have developed more realistic inventory models by relaxing these assumptions in various ways. Let us look at the assumption of constant

holding cost. This assumption presumes that stock can be stored indefinitely to meet future demand. While this is true for many items, certain types of inventories can undergo changes in composition, potency, etc. which make them unfit for consumption. These inventories are, in fact, perishable. To account for this fact, researchers have developed a class of models that treat holding cost as a function of the amount of time that an item is held in inventory. Such models are known as perishable inventory models.

Nahmias (1982) presents a review of the early work in perishable inventory theory. He classifies perishability into two classes: fixed lifetime and random lifetime. The fixed classification includes those cases where the lifetime is known *a priori* to be a specified number of periods or length of time. Of greater interest to the present research is the random lifetime classification which includes those cases where the product lifetime is a random variable with a specified probability distribution. A special case in this classification is that where the lifetime of the product exhibits exponential decay. Nahmias reviews the work of a number of authors who have assumed deterministic demand along with exponential decay of items. In these cases the solution of an ordinary differential equation results in a relationship for inventory position over time which in turn is used to determine the optimal order quantity and optimal total cost. The case of stochastic demand presents a more difficult problem. In general, exponential decay problems with random demand are extremely difficult to handle when there is positive ordering lead time.

Weiss (1982) considers a product facing a constant demand rate, constant replenishment lead-time, a fixed ordering cost, and holding cost that follows a relationship that is nonlinear in time. Specifically the holding cost function is $H(t)=C_h t^{\gamma}$, where C_h and $\gamma \geq 1$ are constants. The author develops an Economic Order Quantity (EOQ) model that minimizes average combined

ordering and holding costs over an infinite horizon. Ferguson et al (2007) show that the EOQ model developed by Weiss (1982) is an approximation of the optimal order quantity for perishable goods such as produce and dairy products sold in small supermarkets that do not receive daily deliveries. Beyond this they apply the Weiss' model to data obtained from a national U.S. grocery chain. Using regression these data are fit to the nonlinear holding cost model. A series of numerical experiments are performed to test the performance of the model against the classical EOQ. The results of the study show Weiss' model delivers superior results under a series of robust scenarios. The improvement over the classical EOQ model is more significant for higher daily demand rate, lower holding cost, shorter lifetime, and a markdown policy with steeper discounts.

The results of the Ferguson et al study leads one to conjecture that ordering in smaller lot sizes may lead to lower operating costs since the cumulative holding costs increase with time since the product is being marked down as it approaches its spoilage date. One approach that leads to smaller lot sizes is investing in setup cost (time) reduction. This idea has received significant attention in the literature over the last twenty-five years.

Porteus (1985) provided a framework for analyzing investment in setup cost reduction in the standard undiscounted EOQ model. In further work, Porteus (1986) extended the framework to the discounted EOQ model. Billington (1987) modified the classical economic production quantity model with backorders to include the setup cost as a function of capital expense. Specifically, relationships are developed for the cases where setup cost varies exponentially and linearly as a function of capital expense. Paknejad and Affisco (1987) investigate the EPQ model

including backorders for the case where setup cost follows a logarithmic investment function. Paknejad and Nasri (1988) do the same for a power investment function. This body of research produces closed form relationships for optimal lot size, optimal setup costs, and optimal total costs, etc., for specific parameter values such as the interest rate and the percentage decrease in setup cost per investment dollar. van Beek and Putten (1987) relate the concept of setup cost reduction to the notion of manufacturing flexibility. The authors specifically investigate a number of situations including the economic lot size when investment cost is an inverse function of setup cost, the two product one machine problem, the case of supplier lead time reduction in (R,Q) systems, and the case of setup cost reduction in a two level, two stage assembly system. Spence and Porteus (1987) modeled the effect of setup reduction on effective capacity in the multiproduct capacitated EOQ model with stationary demands. Kim and Hayya (1989) examined the effective capacity that results from setup reduction in the EPQ model. Zangwill (1987) provided a parametric algorithm that determines the benefits of setup cost reduction in the dynamic lot size model. Nasri et al (1990) considers investing in setup cost reduction in the EOQ model with non-interchangeable demand, backorders, and stochastic lead time. Closed form relationships for optimal lot size, optimal setup cost, optimal total cost, etc. are derived. Results of numerical examples indicate that significant savings may be realized by investing in reduced setup cost when lead time is assumed to be variable.

In this paper we investigate the impact of efforts aimed at improving flexibility, through setup cost reduction, on the optimal order quantity of the EOQ model with nonlinear holding cost as originally formulated by Weiss (1982) and later applied by Ferguson et al (2007) to inventory management of perishable goods.

2. REVIEW OF BASIC MODEL

The basic model considered in this paper is the classic EOQ with deterministic demand, instantaneous replenishment, constant setup cost, and nonlinear holding cost, developed by Weiss (1982) and applied by Ferguson et al (2007) as an approximation of the optimal order quantity of perishable goods. Assuming that the cumulative holding cost for one unit held during t interval of time is $H(t)=C_h t^{\gamma}$, where C_h and $\gamma \geq 1$ are constants, the average inventory cost per unit time, $AC_{basic}(Q)$, and the resulting optimal lot size, Q^*_{basic}, are given by

$$AC_{basic}(Q)=\frac{DK}{Q}+\frac{Q^{\gamma}C_h}{(\gamma+1)D^{\gamma-1}}, \quad (1)$$

and

$$Q^*_{basic}=\sqrt[\gamma+1]{\left(1+\frac{1}{\gamma}\right)\left(\frac{D^{\gamma}K}{C_h}\right)} \quad (2)$$

where

D = demand per unit time (in units),

K = setup cost per setup,

C_h = holding cost per unit per unit time,

AC(Q) = Average inventory cost per unit time,

Q = lot size per order,

$T=\frac{D}{Q}$ = Cycle time, time between placement and receipt of an order.

Note that equation (2) simply reduces to Wilson's square root formula, $Q_W = \sqrt{\frac{2DK}{C_h}}$, when $\gamma = 1$.

Before proceeding, we can easily find the optimal average inventory cost per unit time, $AC^*_{basic}(Q)$, by substituting (2) into (1) as follows

$$AC^*_{basic}(Q) = \sqrt[\gamma+1]{\left(1+\frac{1}{\gamma}\right)^{\gamma} DK^{\gamma} C_h} \tag{3}$$

Again, the results simply reduce to the corresponding results of the classical EOQ model when $\gamma = 1$.

This formulation assumes that the setup cost, K, is constant. In what follows we extend the above model to the case where K is a function of capital expenditure.

3. THE OPTIMAL FLEXIBILITY IMPROVEMENT MODEL

We now assume that the option of investing to reduce K is available. Thus we consider K to be a decision variable and aim at minimizing the average inventory cost per unit time composed of investment to change K, ordering, and nonlinear holding costs. Specifically, we seek to minimize

$$AC_{imp}(Q,K) = ia_K(K) + AC_{basic}(Q), \tag{4}$$

subject to

$$K \le K_0 \ , \tag{5}$$

where i is the cost of capital, $a_K(K)$ is a convex and strictly decreasing function of K representing the investment cost needed to change the setup cost parameter to the level K, $AC_{basic}(Q)$ is the sum of ordering and holding costs given in (1), and K_0 is the original setup cost parameter.

One convenient approach of solving this optimization problem is to use a sequential approach suggested by Porteus (1986). In this case, we ignore constraint (5), fix K, optimize over Q to obtain $AC^*_{basic}(Q)$, and then optimize over K. That is, we seek to minimize

$$w(K) = ia_K(K) + AC^*_{basic}(Q), \tag{6}$$

where $AC^*_{basic}(Q)$ is given by (3). If the optimal K obtained in this way satisfies the $K \le K_0$ restriction, we are done. Otherwise, no investment in flexibility improvement is made and the results of the EOQ model with nonlinear holding cost of the previous section hold. Ofcourse, the problem may become intractable except for some special cases of $a_K(K)$. The following section treats one of these cases which has received significant attention in the literature.

The Logarithmic Investment Function Case

This particular function is used in previous research by Porteus (1985, 1986), Nasri *et al.* (1990), Paknejad et al (1995), and Paknejad *et al.* (2005) dealing with quality improvement as well as setup cost reduction. In this case the following form of investment function is used

$$a_K(K) = A - B\ln K \qquad for \qquad 0 < K \le K_0\ , \tag{7}$$

where $A = B\ln K_0$, $B = \frac{1}{\Delta}$, and Δ is the percentage decrease in K per dollar increase in a_K. Here our main objective is to minimize the following, which is the result of substituting (7) and (3) into (6)

$$w(K) = i(A - B\ln K) + \sqrt[\gamma+1]{\left(1+\frac{1}{\gamma}\right)^{\gamma} DK^{\gamma}C_h}. \tag{8}$$

THEOREM: If Δ>0, then the following hold:

(a) The optimal setup cost parameter, K^*_{Kimp} is,

$$K_{imp}^{*} = \min\left\{K_0, K_{imp}\right\}, \tag{9}$$

where

$$K_{imp} = \sqrt[\gamma]{\frac{(\gamma+1)(iB)^{\gamma+1}}{\gamma D C_h}}, \tag{10}$$

(b) The resulting optimal lot size, Q_{Kimp}^{*}, is given by

$$Q_{K_{imp}}^{*} = \begin{cases} \sqrt[\gamma]{iB\left(\frac{\gamma+1}{\gamma C_h}\right)D^{\gamma-1}} & if \ \ \mathrm{K_{imp}} \le K_0 \\ \\ \sqrt[\gamma+1]{\left(1+\frac{1}{\gamma}\right)\left(\frac{D^{\gamma}K}{C_h}\right)} & if \ \ \mathrm{K_{imp}} \ge K_0 \end{cases}, \tag{11}$$

(c) The resulting optimal investment plus average inventory cost per unit time, $w^{*}(K_{imp})$, is

$$w^{*}(K_{imp}) = \begin{cases} iB\left(\ln\frac{K_0}{K_{imp}} + 1 + \frac{1}{\gamma}\right) & if \ \ \mathrm{K_{imp}} \le K_0 \\ \\ \sqrt[\gamma+1]{\left(1+\frac{1}{\gamma}\right)^{\gamma} D K_0{}^{\gamma} C_h} & if \ \ \mathrm{K_{imp}} \ge K_0 \end{cases}. \tag{12}$$

It is interesting to note that the equations (9) through (12) simplify to the relationships in Porteus (1985) for the traditional EOQ model.

Proof: a) This part is the direct result of minimizing (8) with respect to K. If K minimizes the $w(K)$ of (8), then it is necessary that $K = K_{imp}$ satisfies

$$\frac{\partial w}{\partial K} = -\frac{iB}{K} + \left(\frac{\gamma D C_h}{(\gamma+1)K}\right)^{\frac{1}{\gamma+1}} = 0. \tag{13}$$

Solving (13) yields (10). To prove that K_{imp} is a relative minimum, it is sufficient to show that $\frac{\partial^2 w}{\partial K^2}$ is strictly positive when $K = K_{imp}$, that is,

$$\frac{\partial^2 w}{\partial K^2} = \frac{iB}{K^2} - \left\{ \frac{\gamma D C_h}{\left[(\gamma+1) K \right]^{\gamma+2}} \right\}^{\frac{1}{\gamma+1}} > 0. \tag{14}$$

After some algebraic manipulations, (14) reduces to

$$\frac{\left[(\gamma+1) iB \right]^{\gamma+1}}{K^{\gamma}} > \frac{\gamma D C_h}{\gamma+1}. \tag{15}$$

Upon finding K_{imp}^{γ} from (10) and substituting it for K^{γ}, following a series of simple algebraic manipulations, inequality (15) reduces to $\gamma+1>\gamma$, which is always true since $\gamma \geq 1$. Hence (8) is minimized at K_{imp} and part (a) holds.

b and c) Given $K_{imp} < K_0$, these parts follow after substituting (10) into the appropriate expressions.

4. NUMERICAL EXAMPLE

Consider a situation where initial setup cost K_0 = \$50/setup, demand D = 100 units, γ=2.25, holding cost factor C_h= 0.10, i=0.10 and B=50. Table 1 presents a comparison of the classical EOQ, Weiss EOQ and Weiss EOQ with improved flexibility as a result of setup cost reduction.

Table 1

Comparison of EOQ Models

Variable	EOQ	Weiss EOQ	Weiss EOQ-Invest
Q(units)	316.23	183.74	86.53
K ($)	50	50	4.33
TC ($)	31.62	39.31	19.46
% TC Savings Over EOQ	-	<24.32>	38.46
%TC Savings Over Weiss EOQ	-	-	50.50

The results presented in Table 1 show that investing in flexibility improvement by reducing setup cost results in a significantly reduced order quantity when compared to the classical EOQ and the Weiss EOQ. Accompanying this reduction in order quantity is a 50.5 percent reduction in total inventory cost when compared to the Weiss EOQ. These results suggest that further investigation of flexibility improvement in order quantity models for perishable inventories is warranted.

5. CONCLUSION

This paper presents an extension of the EOQ model with nonlinear holding cost. Specifically, economic trade-offs associated with efforts directed at improving flexibility, through setup cost (time) reduction are studied. Setup cost is viewed as a decision variable and relationships for economic order quantity and optimal average inventory cost per unit time are given based on a logarithmic investment function. Results of a numerical example indicate that there is promise in this modeling approach for order quantities for perishable inventory items.

This paper was partially funded by the Frank G. Zarb School of Business at Hofstra University.

REFERENCES

Billington, P. J. (1987) The classic economic production quantity model with setup cost as a function of capital expenditure. *Decision Sciences* 18(1), 25-42.

Ferguson, M., Jayaraman, V., & Souza, G.C. (2007) Note: an application of the EOQ model with nonlinear holding cost to inventory management of perishables. *European Journal of Operational Research* 180(1), 1-12.

Kim, S.L. & Hayya, J. (1989) Setup reduction and the increase in production capacity. *Proceedings of the Decision Sciences Institute* 855-857.

Nahmias, S. (1982) Perishable inventory theory: a review. *Operations Research* 30(4), 680-708.

Nasri, F., Affisco, J.F., and Paknejad, M.J. (1990) Setup cost reduction in an inventory model\ with finite-range stochastic lead-time. *International Journal of Production Research* 28(1): 199-212.

Paknejad, J., Nasri, F., Affisco, J.F. (2005) Quality improvement in an inventory model with finite-range stochastic lead time. *Journal of Applied Mathematics & Decision Science* 3:177-189.

Paknejad, J. & Affisco, J.F. (1987) The effect of investment in new technology on optimal batch quantity. *Proceeding of the Northeast Decision Sciences Institute* 118-120.

Paknejad, J. & Nasri, F. (1988) Setup cost reduction in the classical production lot size with backorders model. *Proceeding of the Northeast Decision Sciences Institute* 82-86.

Porteus, E. (1985) Investing in reduced setups in the EOQ model. *Management Science* 31(8), 998-1010.

Porteus, E. (1986) Investing in new parameter values in the discounted EOQ model. *Naval Logistics Research Quarterly* 33, 39-48.

Spence, A.M. & Porteus, E.L.(1987) Setup reduction and increased effective capacity. *Management Science* 33(10), 1291-1301.

van Beek, P. & van Putten, C. (1987) OR contributions to flexibility improvement in production inventory systems. *European Journal of Operational Research* 31, 52-60.

Weiss, H. (1982) Economic order quantity models with nonlinear holding costs. *European Journal of Operational Research* 9(1), 56-60.

Zangwill, W.I. (1987) From EOQ towards ZI. *Management Science* 10, 1209-1223.

ON THE MINIMIZATION OF WORKLOAD BALANCING CRITERIA ON IDENTICAL PARALLEL MACHINES

Anthony Cossari, Dipartimento di Economia e Statistica, Universita della Calabria, 87036 Arcavacata di Rende (CS), Italy. a.cossari@unical.it (39) 0984-492431

Johnny C. Ho, Department of Management and Marketing, Turner College of Business and Computer Science, Columbus State University, Columbus, GA 31907, USA. ho_johnny@columbusstate.edu (706) 507-8087

Giuseppe Paletta, Dipartimento di Economia e Statistica, Universita della Calabria, 87036 Arcavacata di Rende (CS), Italy. g.paletta@unical.it (39) 0984-492428

Alex J. Ruiz-Torres, Departamento de Gerencia, Facultad de Administración de Empresas, Universidad de Puerto Rico – Rio Piedras, San Juan, PR 00931, USA. alex.ruiztorres@uprrp.edu (787) 764-0000

ABSTRACT

This paper studies the problem of scheduling jobs on identical parallel processors to minimize workload balancing criteria. While workload balancing is certainly an important criterion given the need of production systems to use all of their resources efficiently, there is no established measure of performance in the scheduling literature that characterizes total workload balance. In this paper, we propose the normalized standard deviation, the normalized mean deviation, and the normalized mean difference as criteria that measure the balance of the workloads among the machines associated to a schedule. A local search heuristic, which performs multiple exchanges of jobs among machines, is presented.

Keywords: Parallel machines scheduling, workload balancing criteria, measures of dispersion, heuristics.

1. INTRODUCTION

In the classical multiprocessor scheduling problem, a set $J = \{1, ..., j, ..., n\}$ of n simultaneously available independent jobs, with a non-negative processing times $p_j > 0$, $j = 1, ..., n$, are scheduled on a set $M = \{1, ..., i, ..., m\}$ of m identical parallel machines. Each machine can process at most one job at a time, and each job must be processed without interruption by exactly one of the m machines. It is supposed that $n > m \geq 2$ to avoid trivialities. A schedule (solution) is represented by an m-partition $S = \{S_1, ..., S_i, ..., S_m\}$ of the set J, where each S_i represents the subset of jobs assigned to the machine i, $i = 1, ..., m$. For each schedule S, the work-loads of the machines are represented by the m-set $C(S) = \{C(S_1), ..., C(S_i), ..., C(S_m)\}$, where $C(S_i) = \sum_{j \in S_i} p_j$ is the work-load of machine i, $i = 1, ..., m$.

Workload balancing is an important practical criterion given the need of production systems to

efficiently use all of their resources. When the work loads are not balanced, resources will become idle waiting for the next set of jobs, for example, a common setup is required for all the parallel resources. It is therefore highly desirable in many production environments with parallel processing to finish the last assigned job for each resource at the same time. However, there is no "established" measure of performance in the scheduling literature that characterizes workload balance. In this paper, the normalized standard deviation, the normalized mean deviation, and the normalized mean difference are proposed as criteria that measure the balance of the workloads among the machines associated to a schedule *S*.

The paper describes the characteristics of each of the proposed measures of workload balance and the need for normalization when multiple instances will be performed as part of an experiment. It also presents an algorithm aimed at workload balancing based on job exchanges. Experiments were performed to evaluate the performance of the algorithm and to determine differences among the criteria. The experiments demonstrated consistency across the three criteria despite the possibility of difference in schedule evaluation.

The remainder of the paper is organized as follows. In Section 2, we describe the three measures of performance for workload balancing rooted in dispersion measures. Section 3 describes an algorithm to generate schedules with minimal unbalance. Finally, concluding remarks and directions for future research are provided in Section 4.

2. PERFORMANCE CRITERIA AND STATISTICAL MEASURES OF DISPERSION

When scheduling jobs on identical parallel machines, an important practical objective is to find the optimal balance of the workloads among the machines, that is to allocate jobs in the machines in such a way that the resulting workloads are as close to each other as possible. Such workload balancing problem is nothing but the basic problem of number partitioning, where a given set of integers are assigned to a collection of subsets so that the sums of the numbers in each subset are as nearly equal as possible. Obviously, an appropriate performance criterion capable of measuring the extent to which balance is obtained needs to be selected. A natural way to approach this desideratum is to look at a measure of the workload deviations and to minimize it. Currently, a number of ad hoc criteria are available in both multiprocessor scheduling and number partitioning contexts. These criteria have been recently investigated by Cossari et al. [4], who demonstrated that they all are essentially well-known statistical measures of data dispersion, and also that two of them are closely related to each other. We now briefly present these existing criteria, then in Section 2.2 the choice of a performance criterion will be put into the broader perspective of choosing an appropriate dispersion measure from those available in statistical methodology.

2.1 Existing Ad Hoc Criteria

Ho et al. [9] addressed the workload balancing problem in multiprocessor scheduling and proposed to measure balance via the *NSSWD* criterion, which is defined for schedule *S* as

$$W(S) = \sqrt{\frac{1}{\mu} \sum_{i=1,\ldots,m} (C(S_i) - \mu)^2}$$

where $\mu = \left(\frac{1}{m}\sum_{i=1,\ldots,m} C(S_i)\right) = \left(\frac{1}{m}\sum_{i=1,\ldots,n} p_i\right)$ is the average of the $C(S_i)$'s, that is the mean completion time. This criterion stems from the well-established practice of measuring data dispersion in terms of the squared deviations from the mean, and has the additional feature of being suited for the relative assessment of multiple problem instances. Later Cossari et al. (2011) showed that $W(S) = \sqrt{m}\frac{\sigma}{\mu}$, where

$$\sigma = \sqrt{\frac{1}{m}\sum_{i=1,\ldots,m}(C(S_i) - \mu)^2}$$

is the popular standard deviation and σ/μ is the coefficient of variation, which is typically adopted in statistical practice to compare dispersions in different groups of data. Thus, except for the constant $\sqrt{m}$, the *NSSWD* criterion is essentially equivalent to the coefficient of variation, implying that both these indices may be equally used as performance measures. Indeed, a little drawback of $W(S)$ is that it inflates $\frac{\sigma}{\mu}$ in an increasing way as m increases. When a given problem instance is to be considered in practical applications, both m and μ have fixed values and therefore the standard deviation σ may just be used being equivalent to *NSSWD* as well as the coefficient of variation.

In the context of number partitioning, it is customary to obtain partitioning by minimizing the maximum pairwise difference between subset sums $C(S_i)$, namely:

$$\Delta(S) = max_{i=1,\ldots,m}\{C(S_i)\} - min_{i=1,\ldots,m}\{C(S_i)\}.$$

This objective function is just the range, which is a well-known, though poor measure of data variability in that it gives the length of the smallest interval that comprises all the data. Going beyond the usual approach for number partitioning, Alidaee et al. [1] proposed the following criterion:

$$x_0 = \sum_{i=1,\ldots,m-1}\sum_{j=i+1,\ldots,m}\left[C(S_i) - C(S_j)\right]^2$$

which allows to recast the problem as an unconstrained quadratic binary program that can be solved by efficient metaheuristic methods. Cossari et al. [4] showed that $x_0 = m^2\sigma^2$, where σ^2 (square of the standard deviation) is the very popular variance, and hence, apart from a factor m^2, x_0 is essentially equivalent to the variance and to the standard deviation itself. When solving a single instance m is fixed, thus the use of x_0 is unnecessary and σ may again be equally employed. If multiple problem instances are to be examined for purposes of comparison, the use of x_0 should be avoided since it is affected by different values of μ so its results are not comparable to each other. Although the criteria $\Delta(S)$ and x_0 have been devised for number partitioning, it should be emphasized that they both might be equally used in the workload balancing problem for multiprocessor scheduling.

Furthermore, a remarkable, perhaps unexpected mathematical relationship between *NSSWD* and x_0 arises, which gives further insight into these criteria (see Cossari et al. [4] for details). In particular, for a given problem instance, $W(S)$ and x_0 turn out to be proportional to each other, thus they may be used interchangeably for both multiprocessor scheduling and number partitioning, together with their associated algorithms. In this situation, indeed, the standard

deviation σ is an equivalent, more simple criterion, which therefore may definitely be used for both problems. Instead, when various instances need to be investigated, for example in comparative studies, the *NSSWD* criterion, or the equivalent coefficient of variation $\frac{\sigma}{\mu}$, are to be preferred to x_0 as they are unaffected by different values of μ.

In the special case of $m = 2$ machines (subsets), all the aforementioned criteria reduce to some function of $|C(S_1) - C(S_2)|$ (Cossari et al. [4]), that is, the absolute difference between completion times (subset sums) in the two machines (subsets). Obviously, minimizing such a function means minimizing the makespan, thus the workload balancing problem is just a minimum makespan problem.

2.2 Criteria Derived from Dispersion Measures

As evidenced by the preceding subsection, any performance criterion for workload balancing is nothing but a statistical measure of dispersion. Moving away from the practice of generating an ad hoc criterion, the task of choosing an appropriate objective function is now viewed from a global perspective on dispersion measures, arguing that, in principle, any of such measures is eligible for use as a performance criterion. In what follows we give a systematic account of dispersion measures, focusing on the most suitable ones for our purposes (a comprehensive treatment may be found, e.g., in Stuart and Ord [11, pp.52–68]). Clearly, they are treated merely as descriptive indices of variability, thus not considering their inferential properties which are not relevant to our framework.

Dispersion measures may be grouped into three basic categories:

(a) Measures based on the difference between certain representative values of the dataset.
(b) Measures based on the deviations of every datum from some central value.
(c) Measures based on the differences of all the data among themselves.

The range of a dataset is the difference between its greatest and least values, thus belongs to category (a). As remarked before, such measure is the most typical objective function in the context of number partitioning, perhaps because of its simplicity. It may be a not very accurate measure of dispersion though, as it simply gives the distance between the extreme values thus ignoring how the bulk of the dataset varies inside the range. It seems that its use in number partitioning or multiprocessor scheduling problems has a poor justification, and hence some more precise measure of variability should be employed instead. Some other measures fall into group (a), such as the interquartile range and the interdecile range (see Stuart and Ord [11, p.52] for more details). Although such measures can be slightly more informative than the range, they still give a rough assessment of dispersion thus seem to be not useful enough in our framework.

When considering measures of category (b), first some central value need to be specified, for example the mean or the median of the dataset. If the mean is used, then the variance naturally arises as the average of the squares of the deviations from the mean. Certainly, the variance is the dispersion measure par excellence in statistical inference, largely due to its straightforward mathematical tractability in the theory of sampling. By extension, it is usually employed for descriptive analysis as well, but a number of alternatives may be equally used in this case as we will show in the following. A drawback of the variance is that its value is expressed in the

squared unit of the data, making its interpretation hard. This is resolved by simply taking the positive square root of the variance thus obtaining the aforementioned standard deviation σ. As shown earlier, recent proposals of performance criteria for multiprocessor scheduling and number partitioning, namely *NSSWD* and x_0, prove to be some function of σ (and the variance), suggesting that their use is unnecessary for single instance problems and the equivalent standard deviation may be used instead.

A different measure falling into group (b) arises if we consider taking the absolute values of the deviations from the mean rather the squares. This leads to:

$$\delta_1 = \frac{1}{m} \sum_{i=1,\ldots,m} |C(S_i) - \mu|$$

which is known as the mean deviation (from the mean). Clearly, this is a very intuitive measure of dispersion around the mean analogous to the variance (and the standard deviation), which may naturally be used as a descriptive index of variability. It has generally little use in inferential statistics because of the difficulty in handling the absolute value in the sampling theory. We, therefore, recommend its possible use as a performance criterion in our framework in addition to the standard deviation. It is worth noting that, in general, the mean deviation is not greater than the standard deviation, and some sharper inequalities may also occur in some cases (Stuart and Ord [11, p.56]).

Sometimes, the central value from which deviations are compiled is assumed to be the median, that is, the middle value in the ordered sequence of data. In this case, one measure in widespread use is the mean deviation from the median, which is the same as the mean deviation defined above, except that the median is used in place of the mean. Such measure has some attractive features which usually make it preferable to the mean deviation and even to the standard deviation, especially when outliers among the data are present. However, the median is a poor central value for our typical small set of $C(S_i)$'s (completion times or subset sums), thus the mean deviation from the median seems not appropriate for use in number partitioning and multiprocessor scheduling and will not be discussed further.

The main representative of category (c) is the mean difference (without repetition) defined in our context by:

$$\delta_2 = \frac{1}{m(m-1)} \sum_{i=1,\ldots,m} \sum_{j=1,\ldots,m} |C(S_i) - C(S_j)|$$

If compared to indices belonging to group (b), such a measure stems from a completely different viewpoint. In fact, while quantities such as the standard deviation or the mean deviation are intended to measure the workload (subset sum) balance as closeness of the $C(S_i)$'s to the mean value, the mean difference does measure balance as closeness of the $C(S_i)$'s to each other. In particular, the mean difference provided above represents the average of the differences of all the possible pairs of $C(S_i)$ values regardless of their sign, where the divisor $m(m-1)$ is the number of pairs given that each $C(S_i)$ is not taken with itself thus explaining why the resulting index is qualified as the mean difference without repetition. A variant is to consider the mean difference with repetition where each $C(S_i)$ is taken with itself also, but the sum of differences clearly remains unchanged and the only distinction lies in the divisor which is now m^2. Obviously, the

two measures are related to each other, thus the former will be just used here and referred to simply as the mean difference. This index was popularized in the statistical community by Gini [5], even if it was used since the 1870s in some astronomical studies (Stuart and Ord [11, p.58]). Like the mean deviation, it suffers from the same difficulty in the mathematical derivations in sampling theory, but it is a truly attractive descriptive measure for our purposes, mainly because it focuses on dispersion of the $C(S_i)$ among themselves which is the natural object in number partitioning or scheduling problems. Several expressions have been devised to simplify the calculation of the numerator of the mean difference; some of them are reported in Stuart and Ord [11, p.62]. Besides, it may be shown that the mean difference cannot exceed $\sqrt{2}$ times the standard deviation. One might argue that a further measure may be obtained if the pairwise differences are squared rather than taken absolutely (this was done by Alidaee et al [1] with their criterion x_0). Unfortunately, such a measure is nothing but twice the variance, as already noted by Cossari et al. [4], thus it carries no theoretical value.

To summarize, apart from the range whose widespread use in number partitioning is highly questionable, three dispersion measures are regarded as fully adequate to quantify the balance of the $C(S_i)$ among the machines (subsets) and thus recommended as performance criteria. These are the standard deviation (or the variance), the mean deviation and the mean difference, all of which are expressed in the same unit of the data with the exception of the variance. Their use, however, should be restricted to problems where a single given instance is to be resolved. There are situations, in fact, where multiple instances need to be studied, for example when a simulation experiment is produced in order to compare different algorithms. In such cases, each of the suggested criteria has values for the various instances which are not fully comparable, being affected by the μ value arising in the associated instance (see also discussion in Cossari et al. [4]).

There are two possible solutions to this comparability problem. One approach is to consider dividing by μ to obtain the ratios $\frac{\sigma}{\mu}$, $\frac{\delta_1}{\mu}$, and $\frac{\delta_2}{\mu}$, which are pure numbers clearly unaffected by the mean and thus appropriate for purposes of comparison. The first of such relative measures, the coefficient of variation, is the best-known one and is widely used in statistical practice for comparing variability. We recall that it coincides essentially with the *NSSWD* criterion introduced by Ho et al. [9]. Another approach to obtain normalized measures ranging between 0 and 1 is by dividing the standard deviation, the mean deviation, and the mean difference by their maximum value, respectively. Such a maximum is the value attained in the assumption of maximum variability, which occurs for a hypothetical schedule where a single machine (subset) has the entire workload (sum of numbers) while the other $m-1$ machines (subsets) have no workload (sum of numbers) at all. By direct calculation it turns out that the following maxima arise:

$$\max \sigma = \mu\sqrt{m-1}; \qquad \max \delta_1 = 2\mu\frac{m-1}{m}; \qquad \max \delta_2 = 2\mu,$$

which generate three normalized criteria as follows:

$$\sigma^{norm} = \frac{\sigma}{\mu\sqrt{m-1}}; \qquad \delta_1^{norm} = \frac{m\delta_1}{2(m-1)\mu}; \qquad \delta_2^{norm} = \frac{\delta_2}{2\mu}.$$

We observe from these expressions that they may also be interpreted as the normalized variants of the relative measures defined earlier. Obviously, these criteria increase towards 1 as

dispersion of the $C(S_i)$'s increases and, like all the preceding measures, have the ideal value of 0 in case of a perfect balance with all of the $C(S_i)$'s equal to each other and thus equal to the mean μ. Having obtained the schedule (partition) that minimizes σ^{norm}, δ_1^{norm} or δ_2^{norm}, the normalized nature of these measures serves the purpose of suggesting how close is the given optimal schedule (partition) to the ideal schedule (partition) of perfect balance, thus assessing the extent to which balance was obtained. Therefore, these normalized performance criteria are especially recommended for multiple instance problems where comparisons need to be made, but also for single instance problems, and will be used later in our simulation experiment. Among them, the best-known in statistical practice is δ_2^{norm}, which does coincide with the famous Gini's coefficient of concentration, usually denoted by R, arising from a different perspective as a measure of inequality in studies on, e.g., income or wealth (e.g., see Stuart and Ord [11, pp.60–64]). In principle, other indices of inequality might be potentially useful in our context, but this possibility will not be explored further in this paper and is possibly deferred to subsequent work.

Finally, we mention the case of $m = 2$ machines (subsets). Once again, every new criterion proposed in this subsection readily proves to be some function of $|C(S_1) - C(S_2)|$, thus confirming that the problems of number partitioning and workload balancing reduce to the problem of minimum makespan whichever criterion is employed to measure balance.

3. ALGORITHM

In the following, we describe a local search algorithm named *Work-load Balancing Algorithm* (*WBA*(*w*)) for balancing the workloads of the machines with respect to the performance measure *w*. It may be applied with any of the related criteria discussed in Section 2, namely the normalized standard deviation, the normalized mean deviation, and the normalized mean difference, by setting $w = \sigma^{norm}$, $w = \delta_1^{norm}$ and $w = \delta_2^{norm}$, respectively.

WBA(*w*) uses three local search procedures that are performed in the sequence. The first procedure, referred to as *Job Interchange Procedure* (*JIP*(*w*)), interchanges one or two jobs performed by a same machine with one or two jobs performed by another machine, if an advantage has been identified, i.e., w decreases for effect of the interchange.

The second procedure is a slightly modified version of *JIP*(*w*), referred to as *JIP*1(*w*), where machines are sorted in non-increasing order of their workloads. Moreover, if *i* and *j* are the machines interested by the current exchange, the jobs in each of these two machines are sorted in non-increasing order of their processing times.

The third procedure, referred to as *JIP*2(*w*), is a further modification of *JIP*1(*w*), in which the jobs of machine *i* are sorted in non-increasing order of their process times, while those of machine *j* are sorted in non-decreasing order of their process times.

Our heuristic returns the best solution among those obtained by using all these procedures in sequence. Obviously, the optimum may be found before the entire sequence is completed. This certainly happens when $w = 0$. With the aim of leading to the better exploration of the solution space, the current solution is restored to the initial feasible solution, before second and third procedures are run. The following is a general schema of the proposed algorithm.

WBA(*w*) *Algorithm*

Step 1. Consider an initial feasible solution $\tilde{S}$, and compute *w*. Set $S_{best} = \tilde{S}$ and $w_{best} = w$. If $w = 0$ (the current solution is optimal) then go to Step 5.

Step 2. Choose $S = \tilde{S}$ as starting point and perform *JIP*(*w*). If $w_{best} > w$ then set $S_{best} = S$ and $w_{best} = w$. If $w = 0$ then go to Step 5.

Step 3. Choose $S = \tilde{S}$ as starting point. Perform *JIP*1(*w*). If $w_{best} > w$ then set $S_{best} = S$ and $w_{best} = w$. If $w = 0$ then go to Step 5.

Step 4. Choose $S = \tilde{S}$ as starting point. Perform *JIP*2(*w*). If $w_{best} > w$ then set $S_{best} = S$ and $w_{best} = w$. If $w = 0$ then go to Step 5.

Step 5. Return S_{best} and w_{best}.

In local search algorithms, an important decision is the selection of the initial solution, i.e., the initial assignment of the jobs to the machines. In the preliminary experimentation the algorithm ran starting from different solutions, obtained via LPT (Graham [6,7]) and MULTIFIT (Coffman et al. [3]) procedures for $P||C_{max}$ (Graham et al. [8]). In general, LPT showed to be the best choice, and therefore it has been chosen for generating the starting solution.

3.1 Job Interchange Procedure (*JIP*(*w*))

Given a current feasible solution *S* and a minimal w scheduling problem, the procedure *JIP*(*w*), for each couple of machine, iteratively interchanges:

- two jobs, *j* and $k \in S_i$, from the machine *i* with two jobs, *u* and $v \in S_l$, from the machine *l*, if the variation, say $w_{u,v}^{j,k}$, of the criterion *w* identifies an advantage.
- the job, $j \in S_i$, from the machine *i* with two jobs, *u* and $v \in S_l$, from the machine *l*, if the variation, say $w_{u,v}^{j}$, of the criterion *w* identifies an advantage.
- two jobs, *j* and $k \in S_i$, from the machine *i* with the job, $u \in S_l$, from the machine *l*, if the variation, say $w_{u}^{j,k}$, of the criterion *w* identifies an advantage.
- the job, $j \in S_i$, from the machine *i* with the job, $u \in S_l$, from the machine *l*, if the variation, say w_{u}^{j}, of the criterion *w* identifies an advantage.

Obviously, the exchange is advantageous if such variations are negative thus resulting in a reduction in the *w* criterion. If the interchange has been carried out then it is necessary to update the current feasible solution.

The variations $w_{u,v}^{j,k}$, $w_{u,v}^{j}$, $w_{u}^{j,k}$ and w_{u}^{j} are determined by setting $w = \sigma^{norm}$, $w = \delta_1^{norm}$ and $w = \delta_2^{norm}$. Formally, the procedure can be described as follows.

JIP(w) Procedure

Step 0. Consider the current feasible solution $S = \{S_1, ..., S_m\}$ and the corresponding $C(S) = C(S_1), ..., C(S_i), ..., C(S_m)$. Set $i = 1$ and $l = m$.

Step 1. $j \in S_i$ and $u \in S_l$

$v \in S_i$ and $v \neq u$

$k \in S_i$ and $k \neq j$

If ($w_{u,v}^{j,k} < 0$) then update the current feasible solution by setting $S_i = S_i \setminus \{j,k\} \cup \{u,v\}$, $S_l = S_l \setminus \{u,v\} \cup \{j,k\}$, $C(S_i) = C(S_i) - p_j - p_k + p_u + p_v$, $C(S_l) = C(S_l) - p_u - p_v + p_j + p_k$. If $w = 0$ then go to Step 3 else set $i = 1$ and $l = m$, and go to Step 1;

End k.

If ($w_{u,v}^{j} < 0$) then update the current feasible solution by setting $S_i = S_i \setminus \{j\} \cup \{u,v\}$, $S_l = S_l \setminus \{u,v\} \cup \{j\}$, $C(S_i) = C(S_i) - p_j + p_u + p_v$, $C(S_l) = C(S_l) - p_u - p_v + p_j$. If $w = 0$ then go to Step 3 else set $i = 1$ and $l = m$, and go to Step 1;

End v.

$k \in S_i$ and $k \neq j$

If ($w_{u}^{j,k} < 0$) then update the current feasible solution by setting $S_i = S_i \setminus \{j,k\} \cup \{u\}$, $S_l = S_l \setminus \{u\} \cup \{j,k\}$, $C(S_i) = C(S_i) - p_j - p_k + p_u$, $C(S_l) = C(S_l) - p_u + p_j + p_k$. If $w = 0$ then go to Step 3 else set $i = 1$ and $l = m$, and go to Step 1;

End k.

If($w_{u}^{j} < 0$) then update the current feasible solution by setting $S_i = S_i \setminus \{j\} \cup \{u\}$, $S_l = S_l \setminus \{u\} \cup \{j\}$, $C(S_i) = C(S_i) - p_j + p_u$, $C(S_l) = C(S_l) - p_u + p_j$. If $w = 0$ then go to Step 3 else set $i = 1$ and $l = m$, and go to Step 1;

End u and j.

Step 2. If $l > i$ set $l = l - 1$ and go to Step 1, otherwise if $i < m - 1$ set $i = i + 1$ and $l = m$ and go to Step 1, otherwise go to Step 3.

Step 3. Return S, $C(S)$ and w.

With regard to efficiency, we note that the procedure considers the interchange of one or two jobs performed by a same machine with one or two jobs performed by a different machine. According to Hubscher and Glover [10] and Anderson et al. [2], the maximum possible neighborhood size occurs when the jobs are evenly divided between the machines, then there are $O(n^4/m^2)$ possible choices to exchange the jobs. In fact, for each of the $m(m-1)$ couples of machines (i, l), we have at most $\lceil n/m \rceil^4$ couples of two jobs (j, k) and (u, v). Note that Step 1 can be performed more times before $JIP(w)$ stops.

The $JIP1(w)$ and $JIP2(w)$ procedures, that are slightly modified versions of $JIP(w)$, attempt to lead the research process to regions of the solution space that have not been previously visited. Formally, the procedures can be described as follows.

JIP1(w) Procedure

$JIP1(w)$ is obtained from $JIP(w)$ by adding at the begin of Step 1: Sort the machines in non-increasing order with respect to their workloads and sort the elementary jobs belonging to each of the subsets S_i and S_l in non-increasing order with respect to their processing times.

JIP2(w) Procedure

$JIP2(w)$ is obtained from $JIP(w)$ by adding by adding at the begin of Step 1: Sort the machines in non-increasing order with respect to their workloads and sort the jobs belonging to S_i in non-increasing order with respect to their processing times and those belonging to S_l in non-decreasing order with respect to their processing times.

4. CONCLUSIONS

Minimizing workload imbalance across parallel resources is a goal of many production environments, yet there is no single measure of performance that has been accepted to characterize this practical goal. In this paper, three measures are described based on the statistical concept of dispersion. The three proposed criteria are described, including the need to normalize as a means to allow for experiments with diverse problem characterization (that is, number of machines, number of jobs, total processing load) to be analyzed properly. The paper presented an algorithm which uses several job exchange rules to maximize work balance. A logical extension of this paper is to test the proposed algorithm by simulation. Other areas of future research include the consideration of these three criteria for related parallel machine problems, particularly those with setups and batching machines.

REFERENCES

[1] Alidaee, B., Glover, F., Kochenberger, G.A. & Rego, C. *A New Modeling and Solution Approach for the Number Partitioning Problem*. Journal of Applied Mathematics and Decision Sciences, 2005, 9(2), 113–121.

[2] Anderson, E.J., Glass, C.A. & Potts, C.N. Machine Scheduling. In: Aarts, E., Lenstra, J.K. (eds.) Local Search in Combinatorial Optimization, pp.361–414. Chichester: Wiley, 1997.

[3] Coffman Jr., E.G., Garey, M.R. & Johnson, D.S. *An Application of Bin-Packing to Multiprocessor Scheduling*. SIAM Journal of Computing, 1978, 7, 1–17.

[4] Cossari A., Ho J.C., Paletta G. & Ruiz-Torres A.J. *A New Heuristic for Workload Balancing on Identical Parallel Machines and a Statistical Perspective on the Workload Balancing Criteria*. Computers and Operations Research, 2011, 39, 1382–1393.

[5] Gini C. Variabilit`a e mutabilit`a. Tipografia di Paolo Cuppini, Bologna (in italian), 1912.

[6] Graham, R.L. *Bounds for Certain Multiprocessing Anomalies*. Bell System Technical Journal, 1966, 45, 1563–1581.

[7] Graham, R.L. *Bounds on Multiprocessing Timing Anomalies*. SIAM Journal Applied Mathematics, 1969, 17, 416–429.

[8] Graham, R.L., Lawler, E.L., Lenstra, J.K. & Rinnooy Kan, A.H.G. *Optimization and Approximation in Deterministic Sequencing and Scheduling: A Survey*. Annals of Discrete Mathematics, 1979, 5, 287–326.

[9] Ho J.C., Tseng T.L.B., Ruiz-Torres A.J. & Lopez F.J. *Minimizing the Normalized Sum of Square for Workload Deviations on m Parallel Processors*. Computers and Industrial Engineering, 2009, 56, 186–192.

[10] Hubscher, R. & Glover, F. *Applying Tabu Search with Influential Diversification to Multiprocessor Scheduling*. Computers and Operations Research, 1994, 21, 877–884.

[11] Stuart, A. & Ord, K. Kendall's Advanced Theory of Statistics, sixth edition. Volume 1: Distribution Theory. Arnold, 1994.

SIDE - A DECISION SUPPORT SYSTEM USING A COMBINATION OF SWARM INTELLIGENCE AND DATA ENVELOPMENT ANALYSIS

Rashmi Malhotra, St. Joseph's University, (610) 660-3497, rmalhotr@sju.edu

ABSTRACT

This study proposes the design of a multi-attribute-decision-support-system that combines the analytical power of two different tools: data envelopment analysis (DEA) and particle swarm optimization (PSO), one of the major algorithms using swarm intelligence. DEA evaluates and measures the relative efficiency of decision making units that use multiple inputs and outputs to provide non-objective measures without making any specific assumptions about data. On the other hand PSO's main strength lies in exploring the entire search space. This study proposes a modeling technique that jointly uses the two techniques to benefit from the two methodologies.

Keywords: Swarm Intelligence, Particle swarm Optimization, Data Envelopment Analysis

INTRODUCTION

A business organization's objective is to make better decisions at all levels of the firm to improve performance. Typically organizations are multi-faceted and complex systems that use uncertain information. Therefore, making quality decisions to improve organizational performance is a daunting task. Organizations use decision support systems that apply different business intelligence techniques such as statistical models, scoring models, neural networks, expert systems, neuro-fuzzy systems, case-based systems, or simply rules that have been developed through experience. Managers need a decision-making approach that is robust, competent, effective, efficient, and integrative to handle the multi-dimensional organizational entities. The decision maker deals with multiple players in an organization such as products, customers, competitors, location, geographic structure, scope, internal organization, and cultural dimension (Porter, 1980). Sound decisions include two important concepts: efficiency (return on invested resources) and effectiveness (reaching predetermined goals). However, quite frequently, the decision maker cannot simultaneously handle data from different sources. Hence, we recommend that managers analyze different aspects of data from multiple sources separately and integrate the results of the analysis. This study proposes the design of a multi-attribute-decision-support-system that combines the analytical power of two different tools: data envelopment analysis (DEA) and particle swarm optimization (PSO), one of the major algorithms using swarm intelligence. DEA evaluates and measures the relative efficiency of decision making units that use multiple inputs and outputs to provide non-objective measures without making any specific assumptions about data. On the other hand PSO's main strength lies in exploring the entire search space. This study proposes a modeling technique that jointly uses the two techniques to benefit from the two methodologies. A major advantage of the DEA approach is that it clearly identifies the important factors contributing to the success of a decision. In addition, I also propose the use of a particle swarm optimization to assess the global minima or maxima that can aid the decision-maker in making decisions regarding the

implications of a decision. One of the important characteristics of population-based search algorithms is their ability to improve the exploration of the search space without falling in the pitfalls of local minima and maxima. The objective of this study is to find a minima solution (ideal loan) nearest to a given loan application using a limited number of iterations. However, not all solutions found using PSO may be true minima so we use the DEA algorithm, a benchmarking technique, to assess the worthiness of the closest optimal solution. The DEA model benchmarks the performance value of the loan application (both original and optimized) against a set of previous loans. Thus, a decision maker can easily analyze and understand any decision using the power of the efficiency frontier algorithm (DEA) and global search algorithm (PSO) to analyze the credit-worthiness of a loan application. The rest of the paper is organized along the following lines. In section II, we provide a review of previous studies on loan evaluation. Section III discusses the model that we use in this study. Section IV discusses the data, methodology, and empirical analysis of our results. Section V summarizes and concludes our study.

LITERATURE REVIEW

Swarm intelligence literature:

Swarm intelligence, based on collective artificial intelligence, is an emerging area in the field of optimization. Many researchers have developed algorithms that model the behavior of different swarm of animals and insects such as ants, termites, bees, birds, fishes, and elephants. In 1990s, researchers introduced two important algorithms – ant colony optimization (Dorgio et. al., 1991) and particle swarm optimization (Kennedy & Eberhart, 1995) based on fish schooling and bird flocking. PSO mimics the movement of birds in a flock sharing information with each other and the way they interact with each other (Acan and Gunay, 2005) defined by topology. The birds in the swarm represent parameter samples called particles. Birds fly around randomly, but keep an eye on others to follow the bird closest to food. Similarly, each particle in the swarm keeps track of its own best solution found so far and shares the information with topological neighbors to fly towards optimal solutions (Brits et. al., 2007). These algorithms have inspired many researchers to create new versions to solve problems in different areas. Researchers have used these models to solve difficult real world problems such as traffic routing, networking, games, industry, robotics, economics, and design of artificial self-organized distributed problem-solving devices. One of the challenges in PSO is to locate global optima without getting trapped in local optima (Hendtlass, 2003). Researchers suggest different topologies to improve the success rate of finding global optima (Brits et. al., 2007). Eberhart et. al. (1996) define gbest and lbest topologies for the original PSO. On the other hand, Kennedy and Mendes (2002) illustrate von Neumann topology, and Suganthan (1999) proposes spatial topology. Similarly, Brits et. al. (2002a) NichePSO, a multi-modal optimization algorithm that employs the Guranteed Convergence PSO (GCPSO) (van der Bergh and Engelbercht, 2002) to improve local convergence while maintaining the diversity of particles by forming subswarms. Li (2004) introduced the Species-based algorithm (SPSO) to alleviate the shortcomings of the k-means algorithm. Likewise, over the last decades, the intelligent behavior of bee swarm have inspired researchers to develop new algorithms. Abbas (2001) proposed a new swarm intelligence technique, Honey Bee Mating Optimization (HBMO) based on marriage in honey bees. This model simulates the evolution of honey-bees starting with a solitary colony (single queen without a family) to the emergence of an eusocial colony (one or more queens with a family). Abbas applied the model to a fifty propositional satisfiability problem (SAT) with 50 variables and 215 constraints and 3-SAT problem (Abbas 2001a, b). Since then, researchers have proposed many variations of bee swarm intelligence. Jung (2003) proposed an evolution method called queen-bee

evolution simulating the queen-bee role in reproduction process. Lu and Zhou (2008b) developed Bee Collective Pollen Algorithm by simulating the honeybees' collective pollen as a global convergence searching algorithm. Researchers have used these algorithms to solve the traveling sales person problem using heuristical approach such as the ant colony optimization (Dorgio & Gambardella, 1997), honey bee mating optimization (Marinakis & marnakis, 2009), and particle swarm optimization (Pang et. al., 2004; Lope & Coelho, 2005; Shi et. al., 2007).

The basic PSO and its variation has also been used to solve problems such as bin-packing problem (Liu et. al., 2008) and flowshop sequencing (Tasgetiren et. al., 2007; Tseng & Liao, 2008). Likewise, researchers have also used ant optimization algorithm to develop a hybrid routing protocol such as works by Rajgopalan & Shen, 2006, Baras & Mehta, 2003, and Camara & Loureiro, 2000) among others. Similarly, HBMO has been successfully applied to solve problems such as partitioning and design of embedded systems (Koudil et. al., 2007) and cluster analysis (Fathian, et. al., 2007), just to name a few. These algorithms have found many industrial applications. For example, US manufacturer of Americal Air Liquids, has achieved significant financial savings by using a computer model based on algorithms inspired by the foraging behavior of ants to analyze every permutation of plant scheduling, gas prices, weather and truck movements. Other applications include telecoms data routing and delivery vehicle fleet scheduling (Bogue, 2008). It is only recently, researchers have started using swarm intelligence to financial service industry. Nenortaite & Simutis (2006) use an intelligent decision-making model to calculate one day forward decision for purchase of the stocks. The decision model is based on the application of artificial neural networks and swarm intelligence technology, PSO. Capital One, the giant financial-services company, has replaced its rigid command-and-control management style with a more flexible approach better suited to a fast-growing business. Like a medium-size ant colony whose territory is invaded by a larger competitor, Capital One constantly searches out and targets new market opportunities. To encourage employees to look for opportunities outside their immediate departments, Capital One revamped its employee evaluation system to reward people who actively search for such "food sources." (Freidman, Catherine, 2003) Martens, et. al., (2010) illustrate the use of Ant Miner+ to build internal rating systems for credit risk. They use Ant Miner+ to infer a propositional rule set from a given data set, hereby using the principles from Ant Colony Optimization.
As illustrated above, while swarm intelligence has been investigated vastly by the optimization community, not many studies have applied this new technique to financial services industry.

Data envelopment analysis literature:

Recently, many studies have illustrated the use of DEA, a non-parametric methodology to analyze different aspects of business entities. The details of the DEA model are discussed in the next section. In contrast to other methodologies, DEA is one of the methods that have traditionally been used to assess the comparative efficiency of homogenous operating units such as schools, hospitals, utility companies, sales outlets, prisons, and military operations. More recently, it has been applied to banks (Haslem, Scheraga, & Bedingfield, 1999) and mutual funds (Haslem & Scheraga, 2003; Galagedera & Silvapulle, 2002; McMullen & Strong, 1998; Murthi, Choi, & Desai, 1997).

Murthi, Choi, & Desai (1997) examine the market efficiency of the mutual fund industry by different investment objectives. They use a benefit/cost non-parametric analysis where a relationship between return (benefit) and expense ratio, turnover, risk, and loads (cost) is established. They also develop a measure of performance of mutual funds that has a number of advantages over traditional indices. The DEA portfolio efficiency index (DEPI) does not require specification of a benchmark, but incorporates

transaction costs. The most important advantage of DEA method as compared to other measures of fund performance is that DEA identifies the variables leading to inefficiencies and the levels by which they should be changed to restore the fund to its optimum level of efficiency. McMullen and Strong (1998) applied DEA to evaluate the relative performance of 135 US common stock funds using one, three, and five-year annualized returns, standard deviation of returns, sales charge, minimum initial investment, and expense ratio. They illustrate that DEA can assist in selecting mutual funds for an investor with a multifactor utility function. The DEA selects optimum combinations of investment characteristics, even when the desired characteristics are other than the two-factors specified in Capital Market Theory. The DEA enable the user to determine the most desirable alternatives, and pinpoint the inefficiencies in a DEA-inefficient alternative. Sedzro and Sardano (1999) analyzed 58 US equity funds in Canada using DEA with annual return, expense ratio, minimum initial investment and a proxy for risk as factors associated with fund performance. Further, they also find a strong relationship among the efficiency rankings using DEA, Sharpe ratios, and Morningstar data. Galagedera and Silvapulle (2002) use DEA to measure the relative efficiency of 257 Australian mutual funds. The further investigate the sensitivity of DEA efficiency to various input-output variable combinations. They find that more funds are efficient when DEA captures a fund's long-term growth and income distribution than a shorter time horizon. In general, the overall technical efficiency and the scale efficiency are higher for risk-aversive funds with high positive net flow of assets.

Haslem and Scheraga (2003) use DEA to identify efficiencies in the large-cap mutual funds in the 1999 Morningstar 500. They identify the financial variables that differ significantly between efficient and inefficient funds, and determine the nature of the relationships. They use Sharpe index as the DEA output variable. They find that the input/output and profile variables are significantly different between the Morningstar 500 (1999) large-cap mutual funds that are DEA performance-efficient and inefficient. Basso and Funari (2001) propose the use of DEA methodology to evaluate the performance of mutual funds. The proposed DEA performance indexes for mutual funds represent a generalization of various traditional numerical indexes that can take into account several inputs and outputs. They propose two classes of DEA indexes. The first class generalizes the traditional measures of evaluation using different risk indicators and subscription and redemption costs that burden the fund investment. The second class of indexes considers a multiple inputs-outputs structure. Thus, they monitor not only the mean return but also other features such as stochastic dominance and the time lay-out. Morey and Morey (1999) present two basic quadratic programming approaches for identifying those funds that are strictly dominated, regardless of the weightings on different time horizons being considered, relative to their mean returns and risks. They present a novel application of the philosophy of data envelopment analysis that focuses on estimating "radial" contraction/expansion potentials. Furthermore, in contrast to many studies of mutual fund's performance, their approach endogenously determines a custom-tailored benchmark portfolio to which each mutual fund's performance is compared. Feroz, Kim, and Raad (2003) illustrate the use of data envelopment analysis to evaluate the financial performance of oil and gas industry. Edirisinghe and Zhang (2007) develop a data envelopment analysis model to evaluate a firm's financial statements over time in order to determine a relative financial strength indicator that can predict firm's stock price returns.

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. Ozcan 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. Halkos and Salamouis (2004) explore the efficiency of Greek banks with the use of a number of suggested financial efficiency ratios for the time period 1997-1999. They show that data envelopment analysis can be used as either an alternative or complement to ratio analysis for the evaluation of an organization's performance. The study finds that the higher the size of total assets the higher the efficiency. Neal (2004) investigates X-efficiency and productivity change in Australian banking between 1995 and 1999 using data envelopment analysis and Malmquist productivity indexes. It differs from earlier studies by examining efficiency by bank type, and finds that regional banks are less efficient than other bank types. The study concludes that diseconomies of scale set in very early, and hence are not a sufficient basis on which to allow mergers between large banks to proceed. Paradi and Schaffnit (2004) evaluate the performance of the commercial branches of a large Canadian bank using data envelopment analysis. Chen, Sun, and Peng (2005) study the efficiency and productivity growth of commercial banks in Taiwan before and after financial holding corporations' establishment. They employ a data envelopment analysis approach to generate efficiency indices as well as Malmquist productivity growth indices for each bank. Howland and Rowse (2006) assess the efficiency of branches of a major Canadian bank by benchmarking them against the DEA model of American bank branch efficiency. Sufian (2007) uses DEA approach to evaluate trends in the efficiency of the Singapore banking sector. The paper uses DEA approach to distinguish between technical, pure technical and scale efficiencies.

Sanjeev (2007) evaluates the efficiency of the public sector banks operating in India for a period of five years (1997-2001) using DEA. The study also investigates if there is any relationship between the efficiency and size of the banks. The results of the study suggest that no conclusive relationship can be established between the efficiency and size of the banks. Lin, Shu, and Hsiao (2007) study the relative efficiency of management in the Taiwanese banking system through DEA. The goal is to estimate the competitiveness of each bank and managerial efficiency is to show the efficiency variation of each bank through Malmquist index. Bergendahl and Lindblom (2008) develop principles for an evaluation of the efficiency of a savings bank using data envelopment analysis as a method to consider the service orientation of savings banks. They determine the number of Swedish savings banks being "service efficient" as well as the average degree of service efficiency in this industry. Hoon and Chunyan (1994) analyzed the productive efficiency of the railway services in 19 Organization for Economic Cooperation and Development (OECD) countries. They report that railway systems with high dependence on public subsidies are less efficient than similar railways with less dependence on subsidies. Cowie and Riddington (1996) evaluate the efficiency of the European railways through the use of a production frontier approach. Yu and Lin (2008) uses a multi-activity network DEA model to simultaneously estimate passenger and freight technical efficiency, service effectiveness, and technical effectiveness for 20 selected railways for the year 2002. Lozano & Gutierrez (2011) illustrate the slacks-based measure of efficiency of 39 Spanish airports using DEA. Liu & Liu (2010) illustrate the use of DEA in evaluating and ranking the research and redevelopment performance of Taiwan’s government-supported research institutes. Saranga & Moser (2010) develop a comprehensive performance measurement framework using the classical and two-stage Value Chain Data Envelopment Analysis model.

As illustrated above, none of the studies illustrate the merger of swarm intelligence and DEA to develop an intelligent decision support system that benefits from the optimal solution finding capability of two diverse techniques. Therefore, the purpose of this study is twofold. Firstly, this study uses DEA to develop a benchmark using actual loan data. The decision support system further uses PSO to find the

global optimal solution beginning with a new loan. The study further validates the PSO solution by benchmarking the optimal loan parameters against the DEA loan base. Thus, the study aims to illustrate the use of PSO systems as a decision-making tool to understand the credit-worthiness of a loan application.

METHODOLOGY

This section illustrates the PSO algorithm and the DEA model. The first part illustrates the models in non-technical terms, and the second part presents the mathematical details of the methodology.

Particle swarm optimization – a swarm intelligence technique:

Swarm is the collective behaviour of decentralized, self-organized systems, natural or artificial. The concept is employed in work on artificial intelligence. The expression was introduced by Gerardo Beni and Jing Wang in 1989, in the context of cellular robotic systems (Beni, et. al., 1989). SI systems typically consist of a set of simple entities co-existing with each other and their environment. These simple agents follow very simplistic rules. Although, there is no centralized control structure that govern how these individual entities should behave. However, they learn from local interactions leading to the emergence of a complex, globalized behavior. A distinguishing characteristics of SI systems is that, by lacking any hierarchical command and control structure, there is no common-mode failure-point or vulnerability (Bogue, 2008). Some natural examples of SI systems inclue foraging ant colonies, honey bee mating, bird flocking, animal herding, bacteria molding, and fish schooling. Particle Swarm Optimization algorithm is a SI technique that imitates humans (or insects) social behavior. Individuals interact with one another while learning from their own experience, and gradually the population members move into better regions of the problem space (Kennedy & Eberhart, 1995). This technique originates from two separate concepts: the idea of swarm intelligence based off the observation of swarming habits by certain kind of animals (such as birds and fish); and the field of evolutionary optimization. In PSO, each particle represents a parameter sample, and the swarm consists of a population of particles. Particles in the swarm share their information with topological neighbors to move around the search space toward optimal solutions.

The data envelopment analysis model:

Data Envelopment Analysis (DEA) (Charnes et al., 1978) model uses linear programming to measure the comparative performance of different organizational units. Further, this generalized optimization technique measures the relative performance of different decision-making entities that have multiple objectives (outputs) and multiple inputs structure. In the DEA terminology, entities/organization units under study are called Decision-Making Units (DMUs). In our study, the DMUs are the loan applications under analysis. DEA measures the efficiency with which a DMU uses the resources available (inputs) to generate a given set of outputs. The DEA methodology assesses the performance of the DMU using the concept of efficiency or productivity, defined as a ratio of total outputs to total inputs. Further, the DEA model estimates relative efficiency, which is with reference to the best performing DMU or DMUs (in case multiple DMUs are most efficient). The DEA allocates an efficiency score of unity or 100 percent to the most efficient unit. The low-performing DMUs' efficiency can vary between 0 and 100 percent in comparison to the best performance.

ILLUSTRATING SIDE MODEL OF THE DECISION SUPPORT SYSTEM FOR LOAN EVALUATION:

To screen consumer loan applications, loan officers use different methods besides intuitive judgment and experience. Using mathematical techniques, many credit-scoring models have been developed to assist the loan officer in differentiating good loans from bad. Besides these traditional statistical models, many financial institutions use artificial intelligence methods, such as expert systems, artificial neural systems, and fuzzy logic. It is only recently that the finance community has started applying data envelopment analysis, a relatively new technique. This study proposes to assess the creditworthiness of a new loan applicant using a decision support system that applies a combination of two diverse analytical techniques: data envelopment analysis and particle swarm optimization. I propose to use the Data Envelopment Analysis methodology to assess the creditworthiness of an existing set of loans whose outcome (accepted and turned good, accepted and turned bad, and reject) is known. The DEA model benchmarks the given set of loans, and assigns an efficiency score.

Each of the loans is a homogenous unit, and we can apply the DEA methodology to assess comparative performance of these loans. The DEA model is a part of a decision support system that uses a number of variables to determine how good a loan is. A loan application includes information such as the applicant's age, housing, address time, total income, number of credit cards, number of dependents, job time, other loan obligations, total debt, monthly rent/mortgage payments, number of inquiries for an applicant, and credit rating. The study creates a DEA model that evaluates the relative efficiency of a set of loans that credit unions have already administered, and allocates a score on the scale of 1 to 100. Further, the next step in the design of the decision support system is to use particle swarm optimization algorithm to map a new loan application to verify how close the loan parameters are to an optimal set of values. Once the PSO model comes up with the optimal loan parameters, the loan officer can discern how close the given loan's variables are to an ideal situation for the given weights on different parameters. To validate the results of PSO, we further benchmark the ideal loan parameters with the given DEA model to check if its 100% efficient, and the PSO algorithm converges to optimal solution.

DATA AND METHODOLOGY

According to Standard & Poor's industry survey, liquidity, inventory, and profit margin are critical to a retailer's success.[1] Therefore, to study the performance of the retail industry (that includes Wal-Mart, Target, Costco, Macys, Sears, J.C. Penney, and BJ Wholesale), we consider seven financial ratios that have been computed on the basis of information contained in the income statement and balance sheet of these firms. The set of ratios that we use to construct the DEA model are: operations (days of sales outstanding/average collection period, inventory turnover, and asset turnover ratios), profitability (operating profit margin, net profit margin, return on equity and return on assets), and financials (quick ratio and total debt/equity ratio). In order to evaluate a firm's financial performance, a financial analyst usually uses these set of ratios.[2] We use the financial statement data available on a quarterly basis from July 2007 to July 2008 from Hoovers Online for this study. Current economic meltdown started in December 2007. Therefore, this time frame allows us to study the financial performance of the retailers before the crisis as well as after onset of economic crisis. Out of these seven ratios, we specify days of

[1] Standard & Poor's Industry Survey, Retailing: General, May 2009.
[2] See Corporate Finance by Ross, Westerfield, and Jaffe, 8th edition, McGraw Hill/Irwin Publishing Company, 2008.

sales outstanding and total debt/equity ratio as input, because for a given company the lower these variables are the better the performance of the company. Similarly, higher operating profit margin, net profit margin, return on equity, return on assets, asset turnover, and inventory turnover imply a better-performing company. Thus, we consider these variables as output variables. Finally, the choice of the DEA model is also an important consideration. We should select the appropriate DEA model with options such as input maximizing or output minimizing, multiplier or envelopment, and constant or variable returns to scale. DEA applications that involve inflexible inputs or not fully under control inputs should use output-based formulations. On the contrary, for an application with outputs that are an outcome of managerial goals, input-based DEA formulations are more appropriate. In addition, for an application that emphasizes inputs and outputs, we should use the multiplier version. Similarly, for an application that considers relations among DMUs, envelopment models are more suitable. Furthermore, the characteristics of the application dictate the use of constant or variable returns to scale. If the performance of DMUs depends heavily on the scale of operation, constant returns to scale (CRS) is more applicable, otherwise variable returns to scale is a more appropriate assumption.

In our study, the comparative evaluation among the companies is an important consideration. Therefore, we select the envelopment models for our analysis. In addition, the outputs are an outcome of managerial goals. Therefore, input-based formulation is recommended for our study. The objective of the analysis is to suggest a benchmark for the seven largest retailing firms. Furthermore, to investigate the effect of scale of operations, if any, among the seven companies, we consider both variable returns to scale and constant returns to scale DEA models. Also, the structure of the DEA model (in envelopment form) uses an equation and separate calculation for every input and output. Therefore, all the input and output variables can be used simultaneously and measured in their own units. In this study, we use the Input-Oriented Variables Return to Scale (VRS) to evaluate the efficiency of seven retail companies. Figure 1 illustrates a decision support system using data envelopment analysis. The decision support system uses the DEA methodology to assess the performance of each company. The DEA-based decision support system uses the company attributes – days of sales outstanding (average Collection Period) and total debt/equity ratio as input variables. The system uses the operating profit margin, quick ratio, return on assets, asset turnover, and inventory turnover as output variables to calculate an efficiency score for a firm. This score is a relative value computed by comparing the given firm to a pool of well-performing companies that serve as a benchmark for the company under evaluation. Each firm is evaluated against the existing firms with an identical set of inputs or outputs that is constructed as a combination of performing and non-performing companies. By using the existing good companies as a "role model," DEA not only helps differentiate well performing (efficient companies from poorly performing (inefficient) firms, but also brings out the reasons why a company may be underperforming.

ILLUSTRATING SIDE MODEL OF THE DECISION SUPPORT SYSTEM FOR LOAN EVALUATION:

To screen consumer loan applications, loan officers use different methods besides intuitive judgment and experience. Using mathematical techniques, many credit-scoring models have been developed to assist the loan officer in differentiating good loans from bad. Besides these traditional statistical models, many financial institutions use artificial intelligence methods, such as expert systems, artificial neural systems, and fuzzy logic. It is only recently that the finance community has started applying data envelopment analysis, a relatively new technique. This study proposes to assess the creditworthiness of a new loan applicant using a decision support system that applies a combination of two diverse analytical techniques: data envelopment analysis and particle swarm optimization. I propose to use the

Data Envelopment Analysis methodology to assess the creditworthiness of an existing set of loans whose outcome (accepted and turned good, accepted and turned bad, and reject) is known. The DEA model benchmarks the given set of loans, and assigns an efficiency score.

Each of the loans is a homogenous unit, and we can apply the DEA methodology to assess comparative performance of these loans. The DEA model is a part of a decision support system that uses a number of variables to determine how good a loan is. A loan application includes information such as the applicant's age, housing, address time, total income, number of credit cards, number of dependents, job time, other loan obligations, total debt, monthly rent/mortgage payments, number of inquiries for an applicant, and credit rating. The study creates a DEA model that evaluates the relative efficiency of a set of loans that credit unions have already administered, and allocates a score on the scale of 1 to 100. Further, the next step in the design of the decision support system is to use particle swarm optimization algorithm to map a new loan application to verify how close the loan parameters are to an optimal set of values. Once the PSO model comes up with the optimal loan parameters, the loan officer can discern how close the given loan's variables are to an ideal situation for the given weights on different parameters. To validate the results of PSO, we further benchmark the ideal loan parameters with the given DEA model to check if its 100% efficient, and the PSO algorithm converges to optimal solution.

Data specifications:

The data for this study is a pooled data set of loans made by nine different credit unions with a total of 185. The applicants can be categorized into three major groups: applicants who were accepted, and were good credits (Group 1); applicants who were accepted, but were not good credits (Group 2); and applicants who applied for a loan, but were rejected (Group 3). Table 1 illustrates the characteristics of the data set. Further, the data set also includes information such as the applicant's age, housing, address time, total income, number of credit cards, number of dependents, job time, co-maker on other loans, total debt, monthly rent/mortgage payments, number of inquiries for an applicant, and credit rating of each applicant. Credit unions in the data set assign loan applicants into four credit groups—excellent (1), good (2), marginal (3), and poor (4). The credit rating is determined on the basis of the number of inquiries. The higher the number of inquiries on an applicant, the lower will be the credit rating. The calculation of credit ratings is consistent across all the credit unions. Thus, based on the information supplied by an applicant, we can calculate the applicant's total payments, total income, and total debt.
The study considers the following variables:

- Total Debt: Total debt of the applicant at the time of application.
- Number of Loans: Total number of loans outstanding in the applicant's name.
- Payments: Total monthly payments
- Dependents: No of dependents of the applicants.
- Total Income: Total monthly income from all sources.
- Job time: Time at current employer.

Data envelopment model specifications for loan evaluation:

The study uses factors such as total debt, number of loans, total payments, number of dependants input variables as they should be minimized. Further, we use total income and time spent in employment as output variables as the higher these are the better it is. Hence, these output variables should be maximized. Besides the mathematical and computational requirements of the DEA model, there are many other factors that affect the specifications of the DEA model. These factors relate to the choice of

the DMUs for a given DEA application, selection of inputs and outputs, choice of DMUs for a given DEA application, choice of a particular DEA model (e.g. CRS, VRS, etc.) for a given application, and choice of an appropriate sensitivity analysis procedure (Ramanathan, 2003). Due to DEA's non parametric nature, there is no clear specification search strategy. However, the results of the analysis depend on the inputs/outputs included in the DEA model. There are two main factors that influence the selection of DMUs – homogeneity and the number of DMUs. To successfully apply the DEA methodology, we should consider homogenous units that perform similar tasks, and accomplish similar objectives. In our study, the loans are homogenous entities as they are from a credit union. Furthermore, the number of DMUs is also an important consideration. The number of DMUs should be reasonable so as to capture high performance units, and sharply identify the relation between inputs and outputs.

In this study, the analysis of a loan emphasizes inputs and outputs. Therefore, I select the multiplier model for my analysis. In addition, factors such as total debt, number of loans, total payments, number of dependants in employment are not very flexible inputs that cannot be immediately controlled. Therefore, output-based formulation is recommended for my study. Furthermore, the quality of the loans does not depend on the scale of operations, thus variable returns to scale is safe assumption. Also, the structure of the DEA model (in multiplier form) uses an equation and separate calculation for every input and output. Therefore, all the input and output variables can be used simultaneously and measured in their own units. Table 2 shows the composition of the data set. There are 23 good loans, 121 loans that were accepted but defaulted and 41 loans that were rejected. Table 3 shows the results of the DEA model to benchmark the loans in the three categories. The DEA model assigns an efficiency score in the range of 1 to 100. Within a tolerance limit of 75%, the DEA model correctly classifies 83% of accepted loans (4 loans less than 75% efficient), 86% bad loans (17 greater than 75% efficient), and 73% reject loans (11 greater than 75% efficient). Thus, DEA model reasonably identifies the good, bad, and reject loans with the probability of type II error very low. However, the cost of type II error, especially for the bad loans is very high. Therefore, we can use a second screening tool to assess the credit-worthiness of a loan. We propose the use of a complementing optimizing tool to examine a loan.

Particle swarm optimization model for the decision support system:

For the PSO model, we use the modified PSO algorithm as explained above. We used a fitness function that minimizes the values of three input variables: the ratio of the total payment to the total income (ratio 1), the ratio of debt to the total income (ratio 2), time in employment, number of outstanding loans, and total number of dependents as the factors that discriminate between a good and a bad loan. We used 43% weight for ratio 1, 42% weight for ratio 2, and 5% weight for time in employment, number of dependents, and total number of loans, respectively. We tested the effectiveness of PSO with two loans, one good and one bad. We used 6-dimensional particles with a swarm size of 40. Each dimension represents the loan parameters: total debt, number of loans, total payments, number of dependants, total income, and time spent in employment. In addition, the swarm parameters are: social component (.5), cognitive component (.5), inertia (.5), and clamping factor (.3). The PSO algorithm found the solution in 5 iterations. Table 4 illustrates the example of a bad loan that was 16% efficient as compared to the existing pool of loans using DEA. However, the optimal solution, found using PSO, is 100% efficient as compared to the existing pool of loans using DEA. Thus, PSO successfully finds the optima. To further test PSO, we used a good loan that was 100% efficient using DEA. The PSO simply leaves the parameter unchanged after 5 iterations indicating that the solution is already optimal. Thus, PSO algorithm successfully discerns good and bad loans that are further validated by data envelopment

analysis algorithm.

SUMMARY AND CONCLUSION

This study proposes the modeling and development of a decision support system that uses a combination of data envelopment analysis and particle swarm optimization. Thus, the decision support system derives benefit from both methodologies to recommend a decision. As illustrated in the literature review section, many studies illustrate the use of PSO and DEA. However, no studies illustrate the fusion of DEA and PSO models. This study illustrates the use of two complementing techniques that can aid a decision maker in making decisions regarding the credit-worthiness of a loan where the cost of accepting a bad loan is very high. DEA does not require the manager to attach prescribed weights to each input and output. Moreover, DEA modeling does not require prescription of the functional forms that are needed in statistical regression approaches. DEA uses techniques such as mathematical programming that can handle a large number of variables and constraints. As DEA does not impose a limit on the number of input and output variables to be used in calculating the desired evaluation measures, it's easier for managers to deal with complex problems and other considerations they are likely to confront. DEA is a methodology based on an interesting application of linear programming allowing a decision maker to use multiple inputs and outputs measured in different units. DEA identifies good units in a given set of DMUs and provides a measure of inefficiency for all others. The DMUs having the most desirable characteristics are rated a score of one (100% efficient), while the DMUs that are inefficient score between zero and one. DEA methodology can identify a bad DMU by comparing its characteristics with a given set of benchmark DMUs having good DMU characteristics.

Similarly, particle swarm optimization does not require restrictive assumptions of the statistical model. PSO model follows the basic principle of swarm intelligence that takes clues from social and cognitive behavior. This model works with costs associated with each parameter. Thus, the decision maker can evaluate the impact of cost on each parameter, something that the DEA model does not associate with the model parameters.

Finally, to illustrate the SIDE model, this study proposed the development of a decision support system to screen consumer loan applications. Loan officers use many different methods besides intuitive judgment and experience. They use mathematical techniques such as credit-scoring models and traditional statistical models. In addition, many financial institutions use artificial intelligence methods such as expert systems, artificial neural systems, and fuzzy logic. This study proposes the development of a decision support system that uses a combination of data envelopment analysis and swarm intelligence. Thus, the decision support system derives benefit from both methodologies to provide a comprehensive review of a loan applicant.

TABLES, FIGURES, & REFERENCES

Tables, figures, references, and full paper available upon request from the authors.

ACCELERATION OF PROGRAM SCHEDULES BY COMBINING TWO FAMILIAR TOOLS: A NUMERICAL EXAMPLE

Jean-Pierre Amor, School of Business Administration, University of San Diego, CA 92110, jpamor@sandiego.edu, 1-619-260-2377

ABSTRACT

Program managers are frequently on the lookout for useful tools and insights that can improve the development of their schedules. This paper illustrates, via a numerical example, the fact that overlapping projects and accounting for the learning phenomenon can have a substantial impact on estimates of the completion times of programs with repetitive projects. To ignore these two factors during the estimation phase of a program is to invite broken delivery promises and the loss of potential contracts.

INTRODUCTION

Effective scheduling is critical to the completion of programs/projects on time and within budget. This paper focuses on the problem of scheduling programs with repetitive projects. These programs deliver a relatively small number of similar manufacturing or construction products, such as aircraft, spacecraft, ships or houses. They also occur in the provision of certain services, such as management consulting, the upgrade of existing equipment, the change of software, or the introduction of a new monitoring system. Frequently, these programs consist of a one-time order for a product that the contractor has never produced before; and because of the complexity of the product, each "unit" requires the execution of a distinct project. Thus, the problems associated with scheduling these programs are very complex; they are also very important, since the cost per unit can be quite high. This effort illustrates the potential acceleration of program schedules by combining two familiar tools: overlapping of projects and accounting for the learning phenomenon.

BACKGROUND

In the scheduling of programs with repetitive projects, a factor that cannot be ignored is the learning curve effect, which suggests that a constant (percent) reduction in a task/activity performance time can be expected with each doubling of the number of repetitions. This relationship is frequently depicted by the log-linear equation:

$$Y=aX^b \quad (1)$$

Where "Y" is the performance time for repetition number "X," "a" is the time/duration of the first repetition and "b" is the ratio of the log of the rate of improvement over the log of 2. The usefulness of this formula in estimating time and cost has been proven over many decades [1].

One of the most commonly used scheduling techniques in project scheduling is the Critical Path Method (CPM). A frequently used convention of CPM is that an activity's duration is given by its standard time: the time a fully experienced resource (e.g., a worker or a crew with its tools and equipment) should take to perform the activity. This convention, however, ignores the fact that, when a less experienced resource is used, the actual activity duration will be longer than under the standard time assumption, even after a fair number of repetitions. This observation led some researchers [2] to demonstrate that, the accuracy of program/project scheduling could be greatly improved by combining learning curves with CPM. Their assumption of "no project overlap," however, was not necessarily realistic, but it was made for the sake of simplicity. Consequently, their potential improvement in scheduling accuracy was probably overestimated.

Although quite promising for improving CPM's accuracy in programs with repetitive projects, the above-mentioned method requires tracking a lot of information. For each repetition of a project, each activity time has to be adjusted for the learning effect, using appropriate parameters in equation (1), before applying CPM. In general a program with N projects, each consisting of M activities, would require MN learning curve calculations and N applications of CPM -- a computationally intensive approach. Although the critical path may change from one repetition to another, the approach does yield, empirically, a "composite learning curve" for the projects in the program.

A project's composite learning curve is not log-linear, but an approximation to that curve, which avoids the computational workload outlined above has been developed [3]. The approximation required: (1) calculating the critical path (CP) times for the first and last project/unit (using appropriate task times calculated from their individual learning curves) and (2) estimating the slope of the log-linear curve passing through these two CP times. The resulting log-linear function, a secant to the actual project composite learning curve, was then used to quickly estimate the duration of all the other projects/units in the program. In several examples, the estimates of overall program duration were correct to within a few percent. In general a program with N projects, each consisting of M activities, would require only 2M learning curve calculations and only two applications of CPM -- regardless of the number of projects in the program. Improved CPM accuracy was now available with a minimal amount of computational effort.

This paper illustrates, via a numerical example, the scheduling of programs with repetitive projects (1) without and with project overlap and (2) without and with accounting for the learning phenomenon. While the "Without Overlap" situation may be applicable in some contexts (e.g., construction of aircraft carriers or space systems), it is not very realistic in other situations (e.g., housing development construction), where several units are usually worked on at the same time. Also, in certain cases, overlapping projects may be an alternative worthy of consideration. Since both learning and overlap affect the unit/project completion times, it seems appropriate to seek insights on the impact of these two factors on the scheduling of programs with repetitive projects.

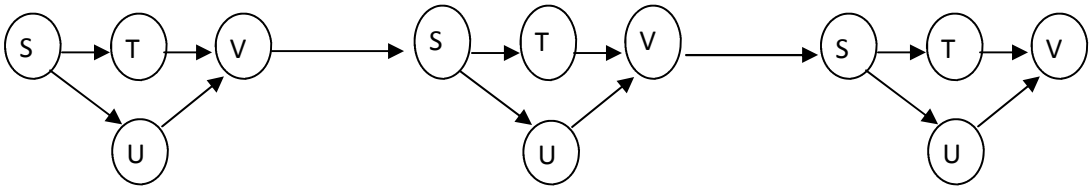
COMBINING TWO FAMILIAR TOOLS: PROJECT OVERLAP AND LEARNING

Since the impact of learning with non-overlapping projects has already been examined [2], we now concentrate on the overlapping situation. However, in order to make meaningful comparisons, we calculate project completion times using the same numerical example for both situations.

Figure 1: Illustration of program connectivity

N = Number of Projects (here N = 3); M = Number of Activities per Project (here M = 4)

a. Without Overlap



b. With (Maximum) Overlap

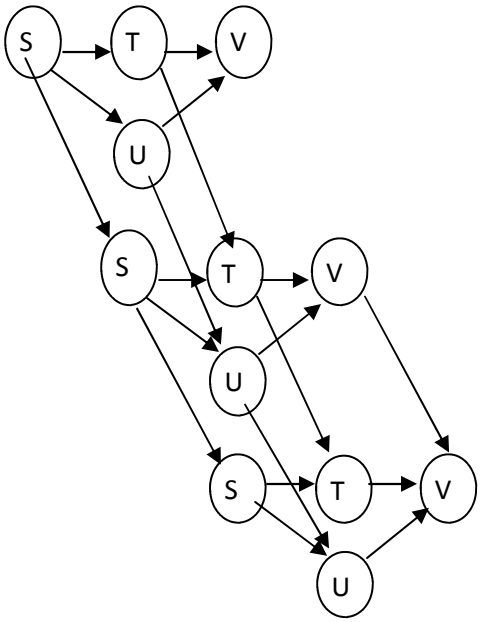


Figure 1 illustrates the connectivity of a simple program network with three projects each consisting of four activities (N=3 and M=4) in two situations: (a) Without Overlap and (b) With (Maximum) Overlap. The first situation might occur when only one large, integrated team is available and must remain together throughout each project. The second situation might obtain when: (1) each activity is performed by one specialized crew, (2) there are as many crews as there are activities, and (3) each crew can move on to the next project (as a single entity) whenever its task on the current project is completed. Of course, partially overlapped projects may occur and may be more likely, but for the sake of simplicity, we limit this discussion to the maximum overlap situation and simply call it "Overlap."

Our numerical example is a slight modification of one used by Amor & Teplitz [3]. It results in more than one CP for some of the projects when they are executed "without overlap and without learning" -- our base case. We examine a program consisting of thirty projects and assume that no crew/team has ever performed its assigned task/tasks on this type of program. This assumption simulates the situation of a one-time order for a product that has never been produced before.

Table 1: Project example

Task/ Activity	Immediate Predecessor(s)	Duration 1st Performance	Rate of Improvement
A (Start)	none	0	n/a
B	A	8.1	.90
C	B	2.8	.95
D	C	11.8	.85
E	D	17.7	.85
F	C	2.0	.90
G	F	21.4	.70
H	F	8.8	.85
I	D	5.9	.85
J	D, G	8.1	.90
K	H, I, J	20.1	.90
L	K	8.8	.85
M	L	2.0	.90
N	L	5.9	.85
O	L	13.2	.80
P	E	5.9	.85
Q	P	2.0	.90
R	C	2.0	.90
S	M, N	4.2	.95
T	O, S	21.4	.70
U	S	4.4	.80
V	Q, R	4.0	.90
W	V	10.1	.90
X (Stop)	T, U, W	0	n/a

Figure 2: AON network for the project example

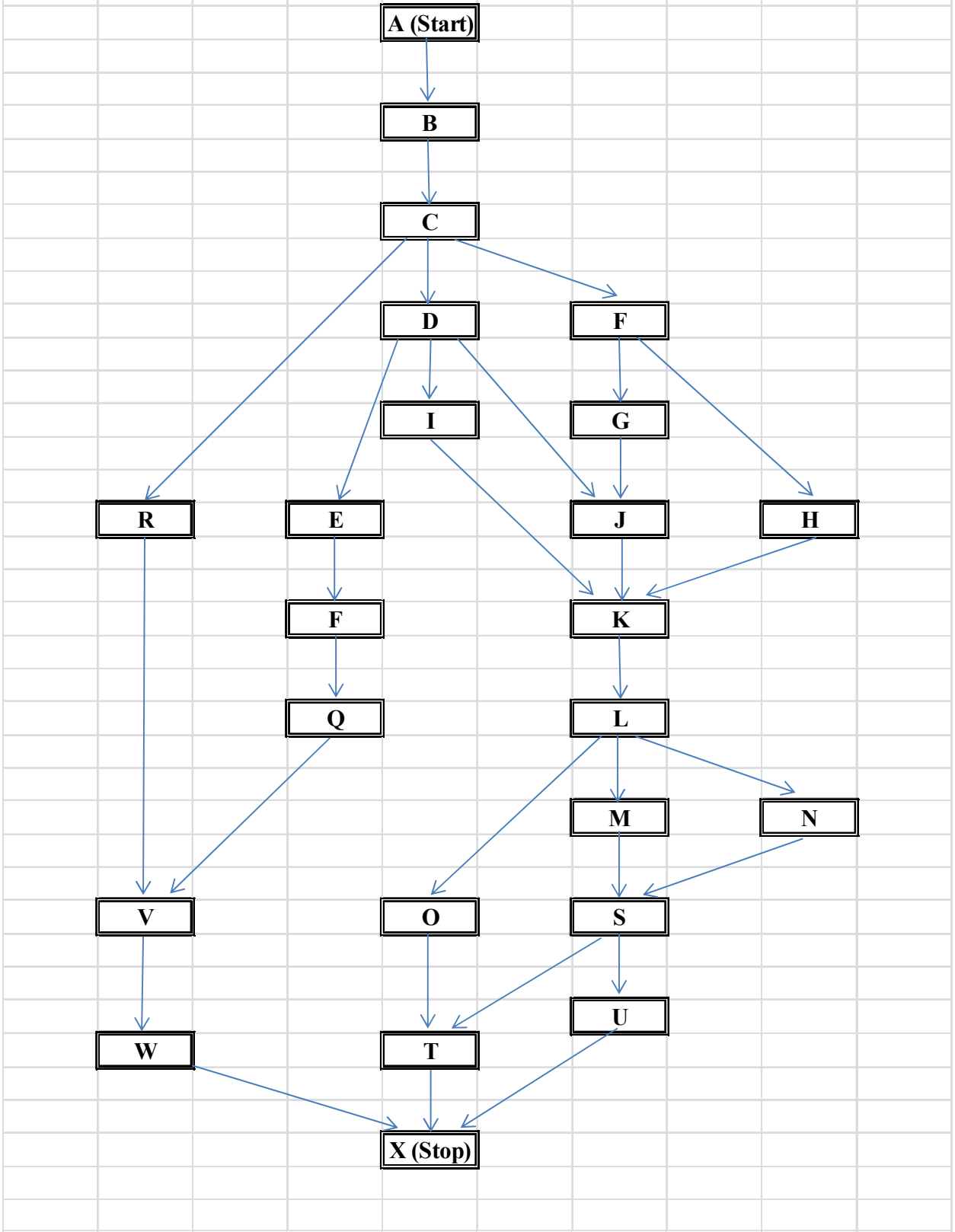


Table 1 lists the tasks/activities, precedence relationships, performance times for the first repetition (without units), and rates of improvement associated with the performance of a project. There are twenty-two individual tasks in the example, which are arranged in a moderately complex network (see Figure 2 for an activity-on-node (AON) network diagram of this project).

Without project overlap

Without project overlap, the critical path (CP) for the first project is A-B-C-F-G-J-K-L-O-T-X. When learning is not taken into account (Case1), the other projects have the same CP and are scheduled for completion 105.9 apart. The overall program duration is estimated at 3,177. When learning is taken into account (Case 2), each activity time is adjusted using appropriate parameters in equation (1). As the individual projects are analyzed, the CP changes three times over the course of the program, and there are two CPs for some of the projects (see Table 2). The scheduled completion time for the overall program becomes 1,626.6. (The individual project/unit completion times for these first two cases are shown later in Table 5.)

Table 2: Project critical paths – without overlap & with learning

Task	Units #1 - #5	Units #6 - #20	Units #21 - #23	Units #24 - #30
A	*	*	* *	*
B	*	*	* *	*
C	*	*	* *	*
D			*	*
E				
F	*	*	*	
G	*	*	*	
H				
I				
J	*	*	* *	*
K	*	*	* *	*
L	*	*	* *	*
M				
N		*	* *	*
O	*			
P				
Q				
R				
S		*	* *	*
T	*	*	* *	*
U				
V				
W				
X	*	*	*	*

An asterisk (*) denotes an activity on the critical path for a given project/unit.
(Note that there are two critical paths for projects/units #21, #22, and #23.)

With project overlap

With project overlap, the CPM calculations are considerably more complex, as suggested by the illustration in Figure 1. For our example, instead of addressing 30 modest networks each consisting of 22 activities, we must face one large network consisting of 660 activities and many more linkages between them.

When learning is not taken into account (Case 3), there are thirty critical paths through the program network. Table 3 illustrates three of them, denoted by ^, *, and + respectively. For example, the 5th critical path (denoted by *) is: A1-B1-C1-F1-G1-G2-G3-G4-G5-J5-K5-L5-O5-T5-T6-T7-T8-T9-T10-T11-T12-T13-T14-T15-T16-T17-T18-T19-T20-T21-T22-T23-T24-T25-T26-T27-T28-T29-T30-X30; where G5 refers to task G on project #5, etc. The multitude of critical paths through the program derives from the fact that activities G and T take exactly the same initial amount of time and learn at the same rate. The overall program duration is now estimated at 726.5. (The times at which each project/unit is completed -- assuming an earliest start schedule -- are listed in Table 5 under Case 3.)

Table 3: Three of the thirty program critical paths for Case 3 -- with overlap & without learning

Unit	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Task																														
A	^*+																													
B	^*+																													
C	^*+																													
D																														
E																														
F	^*+																													
G	^*+	*+	*+	*+	*+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+					
H																														
I																														
J	^				*																				+					
K	^				*																				+					
L	^				*																				+					
M																														
N																														
O	^				*																				+					
P																														
Q																														
R																														
S																														
T	^	^	^	^	^*	^*	^*	^*	^*	^*	^*	^*	^*	^*	^*	^*	^*	^*	^*	^*	^*	^*	^*	^*	^*+	^*+	^*+	^*+	^*+	^*+
U																														
V																														
W																														
X																														^*+

When learning is taken into account (Case 4), each activity time is adjusted using its particular form of equation (1). Table 4 shows the single critical path which occurred through the program network: A1-B1-C1-F1-G1-J1-K1-K2-K3-K4-K5-K6-K7-K8-K9-K10-K11-K12-K13-K14-K15-K16-K17-K18-K19-K20-K21-K22-K23-K24-K25-K26-K27-K28-K29-K30-L30-N30-S30-T30-X30. The overall program duration is now estimated at 473.4. (The times at which each

project/unit is completed -- assuming an earliest start schedule -- are listed at the in Table 5 under Case 4.)

Table 4: Single program critical path for Case 4 -- with overlap & with learning

Unit	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Task																														
A	#																													
B	#																													
C	#																													
D																														
E																														
F	#																													
G	#																													
H																														
I																														
J	#																													
K	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#	#
L																														#
M																														
N																														#
O																														
P																														
Q																														
R																														
S																														#
T																														#
U																														
V																														
W																														
X																														#

RESULTS

Table 5 and Figure 3 present the results for the four cases examined in our example: The "completion time" represents the time at which a project is completed (assuming an earliest start schedule for Cases 3 and 4). The "inter-unit time" for a given project represents the elapsed time between the completion time of the previous project and the completion time of that project. In the "no overlap" cases (Cases 1 and 2), the inter-unit time is simply the time to complete the given project. In the "overlap" cases (Cases 3 and 4), the inter-unit time represents the duration of a subset of project activities.

The shapes of the top two, inter-unit curves, depicted on Figure 3, are as would be expected. Note that the composite learning curve for non-overlapping projects has the familiar, strictly convex appearance; and, although this curve is not log-linear, it is easy to see why it can be closely approximated by a log-linear equation. The bottom two curves, however, are new. They reflect the dramatic improvement (between unit #1 and unit #2) resulting from the overlapping of projects.

When learning is not taken into account (Case 3), no further improvement is indicated. The constant 21.4 difference between the completion times of successive units (see also Table 5)

follows from the fact that the only activity in any given project which does not overlap with the performance of the preceding project is the lengthy activity T.

Table 5: Completion and inter-unit times for all cases

	CASE 1*			CASE 2*			CASE 3*			CASE 4*	
Unit	Compl.	Inter.		Compl.	Inter.		Compl.**	Inter.		Compl.**	Inter.
1	105.9	105.9		105.9	105.9		105.9	105.9		105.9	105.9
2	211.8	105.9		191	85.1		127.3	21.4		120.9	15
3	317.7	105.9		266.4	75.4		148.7	21.4		133.1	12.2
4	423.6	105.9		335.7	69.3		170.1	21.4		143.6	10.5
5	529.5	105.9		400.8	65.1		191.5	21.4		152.9	9.3
6	635.4	105.9		462.7	61.9		212.9	21.4		166.9	14
7	741.3	105.9		522.4	59.7		234.3	21.4		180.9	14
8	847.2	105.9		580	57.6		255.7	21.4		194.6	13.7
9	953.1	105.9		636	56		277.1	21.4		208.4	13.8
10	1059	105.9		690.5	54.5		298.5	21.4		221.9	13.5
11	1164.9	105.9		743.8	53.3		319.9	21.4		235.4	13.5
12	1270.8	105.9		796	52.2		341.3	21.4		248.8	13.4
13	1376.7	105.9		847.2	51.2		362.7	21.4		261.9	13.1
14	1482.6	105.9		897.6	50.4		384.1	21.4		275.2	13.3
15	1588.5	105.9		947	49.4		405.5	21.4		288.1	12.9
16	1694.4	105.9		995.7	48.7		426.9	21.4		300.9	12.8
17	1800.3	105.9		1043.7	48		448.3	21.4		313.7	12.8
18	1906.2	105.9		1091.2	47.5		469.7	21.4		326.5	12.8
19	2012.1	105.9		1138.1	46.9		491.1	21.4		339.2	12.7
20	2118	105.9		1184.6	46.5		512.5	21.4		352	12.8
21	2223.9	105.9		1230.6	46		533.9	21.4		364.3	12.3
22	2329.8	105.9		1276.1	45.5		555.3	21.4		376.8	12.5
23	2435.7	105.9		1321	44.9		576.7	21.4		388.9	12.1
24	2541.6	105.9		1365.7	44.7		598.1	21.4		401.2	12.3
25	2647.5	105.9		1409.9	44.2		619.5	21.4		413.4	12.2
26	2753.4	105.9		1453.8	43.9		640.9	21.4		425.4	12
27	2859.3	105.9		1497.4	43.6		662.3	21.4		437.5	12.1
28	2965.2	105.9		1540.8	43.4		683.7	21.4		449.5	12
29	3071.1	105.9		1583.8	43		705.1	21.4		461.5	12
30	3177	105.9		1626.6	42.8		726.5	21.4		473.4	11.9

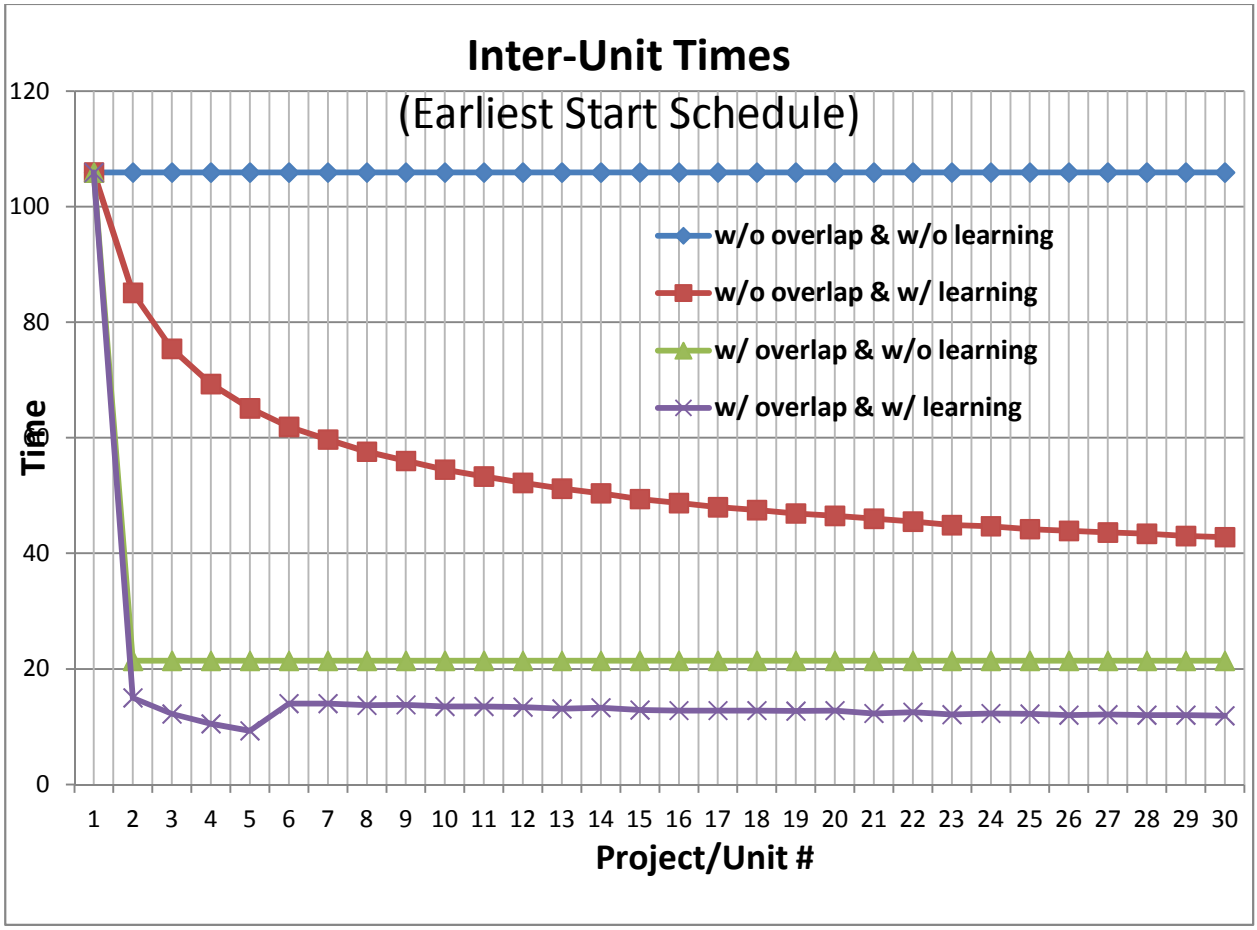
* CASE 1: w/o overlap & w/o learning
CASE 2: w/o overlap & w/ learning
CASE 3: w/ overlap & w/o learning
CASE 4: w/ overlap & w/ learning

** CASES 3 and 4 assume earliest start schedule

When learning is taken into account (Case 4), the inter-unit time curve is not as well behaved. The dramatic drop between unit #1 and unit #2 is due mostly to overlap, but also partly to learning. The portion of the drop from 105.9 to 21.4 is due to overlap. The portion from 21.4 to 15 is due to the learning of the only activity in the second project which does not overlap with the first project (T2). The activity T remains the only non-overlapping activity up to and including unit #5. In other words, the earliest completion time for unit #1 derives from the partial path A1-B1-C1-F1-G1-J1-K1-L1-O1-T1, while that for completing unit #2 results from the partial path A1-B1-C1-F1-G1-J1-K1-L1-O1-T1-T2. This pattern repeats itself up to and

including unit #5, where the partial path is A1-B1-C1-F1-G1-J1-K1-L1-O1-T1-T2-T3-T4-T5. Thus, the inter-unit times continue to decrease based only on the activity T's initial duration and learning rate. The pattern, however, changes at unit 6. At unit #6, the partial path shifts to A1-B1-C1-F1-G1-J1-K1-K2-K3-K4-K5-K6-L6-N6-S6-T6, resulting in the sudden increase in inter-unit time between unit #5 and unit #6. After that, the partial path for unit #7 is the same as that for unit #6 up to K6; then it continues with K7-L7-N7-S7-T7. The partial path for unit #8 is the same as that for unit #7 up to K7; then it continues with K8-L8-N8-S8-T8. This pattern repeats itself until the end of the program at unit #30. The relatively stable pattern from unit #6 onward results in a relatively smooth, generally decreasing inter-unit time curve. The rate of decrease is relatively small since, in going from one project to the next, only five activity changes occur: K_{i+1} is added, and L_{i+1}, N_{i+1}, S_{i+1}, and T_{i+1} replace L_i, N_i, S_i, and T_i respectively; for i=7, …30.

Figure 3: Time between completions of successive projects/units



The partial path shift at unit #6 can be explained in two ways. One way is to see that this shift corresponds to the first change in the project CPs for Case 2 ("without overlap & with learning," see Table 2). In fact, the first five partial paths of Case 4 are based on the first five project CPs of Case 2 (which go from L to O); and the next fifteen partial paths of Case 4 are based on the project CPs for units #6 through #20 of Case 2 (which go from L to N). Another way is to look

at the CPs for each of thirty "with overlap and with learning" nested programs, with N=1, 2, 3, … , 30. When that is done, the thirty (program) CPs associated with these programs correspond exactly to the thirty partial paths associated with Case 4. The growing (program) CPs for the nested programs, as N increases, shift exactly once -- at unit #6.

Table 6 illustrates the impacts of overlap and learning on the scheduled program completion time for our numerical example. Overlap alone tends to reduce the program completion time far more learning alone (77.1% versus 48.8%). In situations of partial overlap, the gap between the two would tend to be smaller than that reported here. As expected, learning tends to have a greater impact on the schedule when there is no overlap (48.8% versus 34.8%). Likewise, overlap tends to have a greater impact on the schedule when learning is not taken into account (77.1% versus 70.9%). And, when both factors are accounted for, the overall effect on the program completion time is quite dramatic -- an 85% reduction. Although these findings are driven by our specific example, the magnitudes involved should leave little doubt as to the potential for acceleration of program schedules by combining project overlap with the learning phenomenon.

Table 6: Acceleration of scheduled program completion times

	Without Learning	With Learning	Percent Reduction
Without Overlap	3177	1626.6	48.8%
With Overlap	726.5	473.4	34.8%
Percent Reduction	77.1%	70.9%	85.1%

CONCLUSION

As organizations increasingly practice project management, useful tools and insights which can improve the development of schedules are increasingly appreciated. This paper has illustrated that, in planning programs with repetitive projects (which are performed in a single sequence) both the learning phenomenon and the overlapping of projects should be accounted for in estimating completion times. And of these two factors, project overlap may well have the greater impact.

Further research in this area might explore: (1) partial, rather than maximum, project overlap and (2) accounting for both overlap and learning in applications that seek an optimal number of parallel sequences of repetitive projects in a given program as discussed in [4]. This should further increase the likelihood of effective scheduling programs with repetitive projects.

REFERENCES

[1] Teplitz, C. J. *The Learning Curve Deskbook*. New York, NY: Quorum Books, 1991.

[2] Teplitz, C. J., & Amor, J. P. Improving CPM's Accuracy Using Learning Curves. *Project Management Journal*, 1993, 24(4), 15-19.

[3] Amor, J. P., & Teplitz, C. J. An Efficient Approximation for Project Composite Learning Curves. *Project Management Journal*, 1998, 29(3), 28-42.

[4] Amor, J. P. Scheduling Programs With Repetitive Projects Using Composite learning Curve Approximations. *Project Management Journal*, 2002, 33(3), 16-29.

CREATING EFFECTIVE PROJECT TEAMS USING PERSONALITY MODELS

Ki-Young Jeong, University of Houston at Clear Lake, Houston, TX 77058, jeongk@uhcl.edu, (281) 283-3862
Ipek Bozkurt, University of Houston at Clear Lake, Houston, TX 77058, Bozkurt@uhcl.edu, (281) 283-3882
Surya T. Sunkara, University of Houston at Clear Lake, Houston, TX 77058, SunkaraS7358@uhcl.edu, (561) 573-9392

ABSTRACT

Human resources and team formation are important issues in any project success. However, very little research has been conducted in project team formation with consideration of diverse characteristics of human resources. Generally, human resources are assigned to project teams based on their availability and technical skills by a project leader. In this research, we considered personalities of human resources as an important factor in creating teams to maximize the team performance and/or to minimize the conflict within teams. For this, three personality models – Myers-Briggs Type Indicator (MBTI), Fundamental Interpersonal Relations Orientation-Behavior (FIRO-B) and Kolbe Conative Index (KCI) – were discussed in the context of project team formation, and we concluded that both FIRO-B and KCI have advantages over MBTI in terms of project team formation since both models provide quantitative performance metrics while the MBTI just indicates types of personality. Finally, we presented a mathematical model to form effective teams using the results from the FIRO-B with a numerical example.

Keywords: Personality Model, Project Team Formation, FIRO-B, KCI.

INTRODUCTION

Forming effective project teams has been an interesting topic for many researchers and practitioners since human resources were considered one of the most important factors in affecting the performance of project teams. Traditionally, human resources have been assigned to project teams based on their availability and technical skills, and the interaction and personality aspects of team members have been often ignored or managed by project manager's experience during this team assignment process.

There is an enormous amount of research on human's personality and its impact on team performance. For example, Peeters et al. [8] investigated the relationship between personality types in the 'Big-Five' model and team performance through meta analysis. According to them, both *agreeableness* and *conscientiousness* have a positive correlation with the team performance. Furnham and Stringfield [3] used the Myers-Briggs Type Indicator (MBTI) to compare the managerial performance in Chinese and European cultures, and they found out that the MBTI generated significant differences between two cultures, however unrelated to robust and multi-factorial measures of management performance. O'Neill and Kline [7] used the 'Big-Five' to analyze the team performance results obtained from a business simulation game, and found out

that both *emotional stability* and an *individual's predisposition* to working in a team setting have a positive correlation with the team performance. LaFollette and Belohlav [5] found out that there was no significant correlation between the team compatibility score computed by the Fundamental Interpersonal Relations Orientation-Behavior (FIRO-B) personality model and the simulated team performance. However, many previous papers did not provide any systematic method to create effective teams based on the results from personality models.

Recently, few researchers started to investigate a systematic way to use personality models to create effective teams. Fitzpatrick and Askin [2] provided a mixed integer programming (MIP) formulation to create effectives teams based on results from the Kolbe Conative Index (KCI) model. André et al. [1] presented a formal model to assign human resources to teams where they used the Delphi technique to propose software project roles and competences, and used psychological tests including the MBTI and data mining tools to form a project team.

The objective of this paper is aligned with those of Fitzpatrick and Askin [2] and André et al. [1]. We discuss three personality models, MBTI, KCI, and FIRO-B to investigate their fitness to systematically create effective teams based on their results. Specifically, we attempt to develop a mathematical model to form effective teams when the results from FIRO-B are given.

The remaining of this paper is organized as follows: Personality models to discuss the MBTI, KCI and FIRO-B in the context of team formation, mathematical formulation to develop effective teams based on the results from personality models, numerical examples to demonstrate the mathematical models, and conclusions and future studies.

PERSONALITY MODELS

In this section, three personality models are briefly explained in the context of the project team formation.

Myers-Briggs Type Index (MBTI)

The MBTI is a personality model developed by Katherine Briggs and Isabel Myers based on Jung's theory of psychological types, and it is the most popular personality type model in US [4]. The MBTI uses the following four dimensions of preferences and each dimension has two opposite preferences:

(1) Orientation of Energy: Extrovert (E) and Introvert (I). Extrovert people obtain energy from outer world while working with teams whereas introvert people obtain it from inside themselves. They may drain energy from interaction in a team.
(2) Manner of Information Gathering: Sensing (S) and Intuitive (N). People who prefer sensing rely on facts and reality, focusing on details while intuitive people rely on intuition, possibilities, and imagination, focusing on the big picture.
(3) Manner of Decision-Making: Thinking (T) and Feeling (F). Thinkers are logical, analytical, and more objective than feelers. They make decisions based on principles, policies, and criteria. Feelers use values, emotion, devotion, and opinions from others. They want to maintain harmony within a team.

(4) Attitude toward Structure: Judging (J) and Perceiving (P). People who prefer judgment are outcome-oriented and decisive, and like closure and task completion. They tend to establish due date and stick to it while people who prefer perceiving are more process-oriented, and prefer to keep things open and flexible. They seek additional information and new possibilities.

Based on two opposite preferences with four dimension, the MBTI generates total sixteen different preference type combinations (e.g. E or I; S or N; T or F; J or P). Hence a person may be identified to one of those sixteen combinations (e.g. ESTJ, ESTP etc). Myers [6] claimed that the best team performance is achieved by team members who differ on one or two preferences to complement each other, and who have two or three common preferences for better communication and understanding. She also recommended that types S and N are useful as well as types T and F. However, she claimed that people who have significant difference on J and P preferences will have difficulty to understand each other.

Although the MBTI is a very useful instrument to form teams, it does not provide any specific quantitative metrics to be used in the team formation process. Therefore, when project managers use the MBTI to select team members, they need to follow Myers's recommendations mentioned above. In other words, it is still difficult to establish a systematic mathematical model to create teams with the MBTI since these recommendations do not completely characterize team formation processes.

Kolbe Conative Index (KCI)

Although the KCI is considered as a personality model in this study, it is slightly different from MBTI and FIRO-B in that it measures conative or individual's instinctive behaviors that appear while that individual tries to achieve goals. The KCI classifies instinctive behaviors into four categories: *probing*, *patterning*, *innovating* and *demonstrating*. Each category has its corresponding 'Action Mode' described below:

(1) Fact Finder (FF): This is an action mode corresponding to the *probing* instinctive behavior. This mode is related to the information gathering. Fact finders are concerned with details, strategies, and research. They collect and analyze data, and establish priorities before making decisions.
(2) Follow Through (FT): This is an action mode corresponding to the *patterning* instinctive behavior. This mode is related to structure. People controlled by FT seek structure, and plan and schedules in advance. They also behave in a sequential manner.
(3) Quick Start (QS): This is an action mode corresponding to the *innovating* instinctive behavior. This mode is related to the way to deal with risk and uncertainty. Quick starters initiate changes, take risk, and they are innovative.
(4) Implementor (IMP): This is an action mode corresponding to the *demonstrating* instinctive behavior. This action mode is related to the way to handle space and tangibles. Implementors are good at using space and materials, develop constructs, and easily use hand-on equipment.

An individual has all of these four instincts to some degree, and can operate in any of these modes. Each of action modes can be classified into the following three operating zones based on the ten-point scale calculated from a set of questions: Prevent, Respond, and Initiate.

(1) Prevent: A prevention-operating zone for an action mode has a *score of 1 to 3*. This zone basically represents the resistance or how the individual won't act.
(2) Respond: A response or accommodating zone for an action mode has a *score of 4 to 6*. This zone represents how an individual is willing to act.
(3) Initiate: An initiating zone for an action mode has a *score of 7 to 10*. This zone represents how an individual will act. In other words, the individual will tend to initiate that type of behavior and is comfortable working in that way.

Let us suppose that each individual has its own KCI score. Then the question becomes: how to create effective teams? According to the KCI, team's success or failure is dependent on the balance of conative energy inside the team. A raw team score is developed by computing the distribution of team members across operating zones for each action mode. Then, the team *synergy* is calculated by computing the distribution of average percent of each zone across action modes. Figure 1 displays the team synergy values from an example in [2]. The ideal synergy in the final column is given by Kolbe Corp. through experimental evidence.

The deviation from the ideal synergy – either positive or negative – is converted into the term *profitability*, a score out of one hundred. A team with lower deviation from ideal team synergy values is considered to have higher synergy, resulting in high *profitability*. In general, a team with higher synergy is considered more successful since the team as a whole entity can use an appropriate amount of energy to initiate solutions, triggering appropriate actions required, and preventing further problems from occurring. For example, there is only one percent positive deviation from the ideal synergy in Initiate zone in Figure 1 while there is no deviation in Prevent and Response zones.

Zone/Action Mode		FF	FT	QS	IMP	Team Synergy (zone average)	Ideal Synergy
Prevent (30%)	1	7%	11%	62%	33%	28%	20-30%
	2						
	3						
Response (60%)	4	14%	53%	20%	67%	41%	40-60%
	5						
	6						
Initiate (30%)	7	69%	36%	18%	0%	31%	20-30%
	8						
	9						
	10						

Figure 1. Kolbe Team Synergy

When there is excessive conative energy in a combination of action modes and operating zones, the *inertia* occurs. According to Kolbe Corp, the ideal conative energy for both Prevent and Implement zones is 30% while it is 60% in Response zone as displayed in the first column of Figure 1. The positive deviations (excess values) of 12 action mode and zone combinations from the ideal conative energy value for zone are added to compute the *inertia*. For example, the positive excess value (inertia) in Figure 1 is 86, given by [(62-30) + (33-30) for Prevent zone] +

[(67-60) for Response zone] + [(69-30) + (36-30) for Initiate zone]. The KCI uses a term called, *goal attainment*, to measure the level of *inertia* in a team. The lower goal attainment means a stagnation of energy in a team. Therefore, it is apparent that we have to consider both team synergy and inertia when creating teams based on KCI.

Fundamental Interpersonal Relations Orientation-Behavior (FIRO-B)

The FIRO-B personality model was originally developed by Schutz [9]. It intends to measure the intensity of the interpersonal needs of *inclusion* (I), *control* (C), and *affection* (A). *Inclusion* refers to the need of individual's social orientation – belongness and interaction. *Control* refers to the need for power and influence, related to leadership behavior. *Affection* indicates the need for intimacy and friendship. Each of these needs has two dimensions to describe how much each of the three needs is *expressed* (*e*) or *wanted* (*w*). The *expressed* refers to the degree to which one behaves in that way toward others while the *wanted* describes the degree to which one wants others to behave that way toward oneself.

The FIRO-B model describes the interaction of the three interpersonal needs with the *expressed* and *wanted* dimensions of each need, resulting in six categories of interpersonal needs as described in Figure 2. Each category is measured with a 10 point scale (0-9).

Dimension/Need	I	C	A
e	The extent of your effort to include others in your activities.	The extent of your effort to control and influence others.	The extent of your effort to get close to people, and to engage them on a personal level.
w	The extent of your wish others to include you in their activities.	The extent of your comfortability working in well-defined situations with clear instructions.	The extent of your wish others to act warmly toward you.

Figure 2. FIRO-B Model

Suppose that we have FIRO-B scores for two persons, *i* and *j*. Then, the following three different types of interpersonal incompatibility can be defined between these two individuals: *Reciprocal incompatibility* (RI_{ij}), *Originator incompatibility* (OI_{ij}), and *Interchange incompatibility* (II_{ij}). *Reciprocal incompatibility* refers to the match between one's need for expressed behavior and the other's need for wanted behavior. This incompatibility occurs when one has a high level of expressed need while the other has a low level of the wanted need – see Equation (1) below, and any score higher than 6 indicates a strong possibility of incompatibility. *Originator incompatibility* indicates the match between two people's expressed needs. This incompatibility occurs when both want to initiate something or when neither wants to do – see Equation (2) below, and any score higher than 6 or lower than -6 indicates a strong possibility of incompatibility. *Interchange incompatibility* refers to the extent to which both have a similar total level of need in an area. For example, this incompatibility occurs when one emphasizes the affection need highly while the other emphasizes the control need highly – see Equation (3) below, and any score higher than 6 indicates a strong possibility of incompatibility.

The formulae representing these three incompatibilities between two persons *i* and *j* are given in Equations (1), (2), and (3).

$$RI_{ij} = |e_i - w_j| + |e_j - w_i| \quad (1)$$

$$OI_{ij} = (e_i - w_i) + (e_j - w_j) \quad (2)$$

$$II_{ij} = |(e_i + w_i) - (e_j + w_j)| \quad (3)$$

Figure 3 summarizes nine possible incompatibility combinations with their desired values. As seen in Figure 3, the row sum provides the overall reciprocal incompatibility (*RI*), overall originator incompatibility (*OI*), and overall interchange incompatibility (*II*) while the column sum provides the overall inclusion incompatibility (*IIC*), overall control incompatibility (*CIC*), and overall affection incompatibility (*AIC*). Finally, it also provides the total incompatibility (*IC*) as sum of all column sums or row sums. Please note that the desired score (threshold value) is displayed in the parenthesis.

Incompatibility Type	**Areas**			**Row Sum**
	I	**C**	**A**	
Reciprocal	RI_{ij} (+6)	RI_{ij} (+6)	RI_{ij} (+6)	Overall *RI*
Originator	OI_{ij} (±6)	OI_{ij}(±6)	OI_{ij}(±6)	Overall *OI*
Interchange	II_{ij}(+6)	II_{ij}(+6)	II_{ij}(+6)	Overall *II*
Column Sum	Overall *IIC*	Overall *CIC*	Overall *AIC*	Total *IC*

Figure 3. Summary of Incompatibility and Areas

Therefore, the summary in Figure 3 suggests a way to create effective teams – create teams such that the resulting total incompatibility is minimized, which will be discussed in the next section.

MATHEMATICAL FORMULATION

Based on the reason previously discussed, the MBTI is excluded from further consideration to form effective teams using quantitative approaches. Hence we discuss mathematical models using KCI and FIRO-B here. Suppose that we need to create multiple teams from a pool of workers with specific skills. Each team requires the specific number of employees with specific skills from the pool.

Team formulation with KCI

Fitzpatrick and Askin [2] developed the MIP model for the situation described above.

i = a subscript for worker;
j = a subscript for skill group;
k = a subscript for team;
z = a subscript for an operating zone – 1 for Prevent; 2 for Response; 3 for Implement;
m = a subscript for an action mode – 1 for FF;…,; 4 for IMP;

$$Y_{ik} = \begin{cases} 1 & \text{if worker } i \text{ is assigned to team } k; \\ 0 & \text{otherwise;} \end{cases}$$

K_j = set of workers in skill group j;

$$a_{izm} = \begin{cases} 1 & \text{if worker } i \text{ exhibits operation} \\ & \text{zone } z \text{ for action mode } m; \\ 0 & \text{otherwise;} \end{cases}$$

S_{jk} = number of workers of skill group j required for team k;

Z_z = desired level of zone z across all modes m ($Z_1 = Z_3 = 1$, and $Z_2 = 2$);
Z_{zm} = maximum desired level of mode m at zone z ($Z_{1m} = Z_{3m} = 30\%$, and $Z_{2m} = 60\%$);
d_{zk}^{+}= negative deviation (deficiency) of team k's synergy from an ideal value for operating zone z. For example, if team k's synergy at Prevent zone is 15%, then d_{1k}^{+}= 5, negative deviation from 20.
d_{zk}^{-}= positive deviation (excess) of team k's synergy from an ideal value for operating zone z. For example in Figure 1, d_{3k}^{-}= 1 for team k since its Initiate synergy value is 31% while its ideal range is 20~30%;
d_{zmk}^{-} = positive deviation (excess) of team k's energy from an ideal value for operation zone z and action mode m combinations.

The labor pool extraction model with multiple teams (LPEMT) is given by

$$\min \sum_k \sum_z (d_{zk}^{+} + d_{zk}^{-}) + \sum_k \sum_z \sum_m d_{zmk}^{-} \tag{4}$$

subject to:

$$\sum_{i \in K_j} Y_{ik} = S_{jk}, \ \forall j, k \tag{5}$$

$$\sum_k Y_{ik} \leq 1, \qquad \forall i \tag{6}$$

$$\sum_m \sum_i a_{imz} Y_{ik} + d_{zk}^{+} - d_{zk}^{-} = Z_z \sum_j S_{jk} \ \forall z, k \tag{7}$$

$$\sum_i a_{imz} Y_{ik} + d_{zmk}^{+} - d_{zmk}^{-} = Z_{zm} \sum_j S_{jk} \forall z, m, k \tag{8}$$

for all variables ≥ 0;

Please note that Y_{ik} is a decision variable to decide which worker is assigned to which team. The first term in objective function (4) computes the deviations of team energy for all zones from ideal KCI values across all teams while the second term calculates the deviation (excess) of energy for all zone-mode combinations across all teams. Hence the objective function computes all deviations from optimal KCI values across teams. Equation (5) ensures that each team has the appropriate number of workers from each skill group. Equation (6) ensures that a worker is assigned to at most one team. Equation (7) defines the deviation for each team from the optimal synergy by adding energy deficiency and by subtracting energy excess to and from the optimal synergy, respectively. Equation (8) defines the deviation for each team from the optimal inertia by adding energy deficiency and by subtracting energy excess to and from the optimal inertia, respectively.

Team Formulation with FIRO-B

Now we consider the FIRO-B based LPEMT model to minimize the total incompatibility (LPEMT-TIC). The following notations are additionally considered.

i, j = subscripts for worker;
l = a subscript for skill;
r = a subscript for need (Inclusion, r = 1; Control, r = 2; Affection, r = 3);

S_{lk} = number of workers of skill group l required for team k;
K_l = set of workers in skill group l;
a_{ir}^{e} = measure of $expressed$ intensity of need r for worker i;
a_{ir}^{w} = measure of $wanted$ intensity of need r for worker i;
RI_{rk} = reciprocal incompatibility of team k for need r;
OI_{rk} = originator incompatibility of team k for need r;
II_{rk} = interchange incompatibility of team k for need r;

Then the LPEMT-TIC is given,

$$\text{Min } \sum_k \sum_r (RI_{rk} + |OI_{rk}| + II_{rk}) \tag{9}$$

Subject to:

$$\sum_k \ Y_{ik} \leq 1, \forall\, i \tag{10}$$

$$\sum_{i \epsilon K_l} Y_{ik} = S_{lk,} \ \forall\, l, k \tag{11}$$

$$\sum_i \sum_{j, j \neq i} Y_{ik} Y_{jk} \left(\left| a_{ir}^{e} - a_{jr}^{w} \right| + \left| a_{jr}^{e} - a_{ir}^{w} \right| \right) = RI_{rk} \quad \forall\, r, k \tag{12}$$

$$\sum_i \sum_{j, j \neq i} Y_{ik} Y_{jk} \left\{ \left(a_{ir}^{e} - a_{ir}^{w} \right) + \left(a_{jr}^{e} - a_{jr}^{w} \right) \right\} = OI_{rk} \quad \forall\, r, k \tag{13}$$

$$\sum_i \sum_{j, j \neq i} Y_{ik} Y_{jk} \left| \left(a_{ir}^{e} + a_{ir}^{w} \right) - \left(a_{jr}^{e} + a_{jr}^{w} \right) \right| = II_{rk} \quad \forall\, r, k \tag{14}$$

for all variables ≥ 0;

The objective function (9) minimizes the total sum of three incompatibilities across all teams. Please note that we consider the absolute value of the interchange incompatibility in the objective function, and Y_{ik} is a decision variable. Again, equation (10) ensures that each team has the appropriate number of workers from each skill group, and equation (11) ensures that a worker is assigned to at most one team. Equations (12), (13), and (14) compute the *RI*, *OI*, and *II* across all teams for each need, respectively. That is, LPEMT-TIC model assigns each worker to each team such that the total incompatibility across teams is minimized.

Please note that LPEMT-TIC has nonlinear constraints and terms, $Y_{ik}Y_{jk}$ for Equations 12 to 14. To avoid any technical complexity involved in solving these nonlinear constraints, we use Excel spreadsheet model with the Excel Premium Solver (EPS).

NUMERICAL EXAMPLES

Two numerical examples are provided: One for LPEMT, the other for LPEMT-TIC. All data are randomly generated.

Team Formulation with KCI Example

The problem descriptions and results are displayed in Figure 4. We consider a pool of eight workers with two different skill groups, and want to create two teams where each team needs two workers from each skill group – Hence four workers each team. The KCI profiles for two different skill groups (skill group 1 – workers 1, 3, 5, and 7; skill group 2 – workers 2, 4, 6, and 8) are provided for each operating zone (P – Prevent; R – Response; I – initiate) for each action mode. For example, a worker 1 has the (1, 0, 0; 0, 0, 1; 1, 0, 0; 0, 0, 1) profile, which may be rewritten to the zone-mode matrix below:

	FF	FT	QS	IMP
P	1	0	1	0
R	0	0	0	0
I	0	1	0	1

Please note that there are 144 (= 3 x 4 x 3 x 4) different combinations of KCI scores equivalent to the values of this matrix – e.g. 3827, 2718 etc where each digit represents KCI score of FF, FT, QS and IMP mode. For example, '3827' is called a modus operandi (MO) score, representing 3 (P) in FF; 8 (I) in FT; 2 (P) in QS; 7 (I) in IMP. The shaded cells 'N56' to 'O59' ('N56:O59') and 'N64' to 'O67' ('N64:O67') display the output valued of decision variables obtained by EPS for skill group 1 and 2, respectively. Please note that '1' indicates that a corresponding worker is assigned to team 1 or 2. The cells 'G75' and 'N75'compute deviations for zone-mode combinations for teams 1 and 2, respectively where its team synergy (TS) is also calculated for each zone across modes. The cell 'L79' is the target objective function to be minimized by EPS, which is sum of cells 'G75' and 'N75'. According to the results in Figure 4, team 1 consists of workers 1, 5 (from skill group 1) and workers 2 and 6 (from skill group 2) while the team 2 has workers 3, and 7 (from skill group 1) and workers 4 and 8 (from skill group 2), and this assignment generates 217 as its total deviation.

Team Formulation with FIRO-B Example

LPEMT-TIC based example for the same situation is displayed in Figure 5. The FIRO profiles for all eight workers are described in the cells 'B5:G8' and cells 'B17:G20' for teams 1 and 2, respectively. Decision variables are calculated in cells 'H5:I8' and cells 'H17:I20' for teams 1 and 2, respectively. Based on the values in the decision variables indicating workers' assignment to each team, three incompatibilities for two teams are calculated and displayed in cells 'C27:F29' for team 1 and cells 'I27:L29' for team 2. Their corresponding total team incompatibility is computed at 'F30' and 'L30' for teams 1 and 2, respectively. The minimized total incompatibility across teams is calculated at E35'.

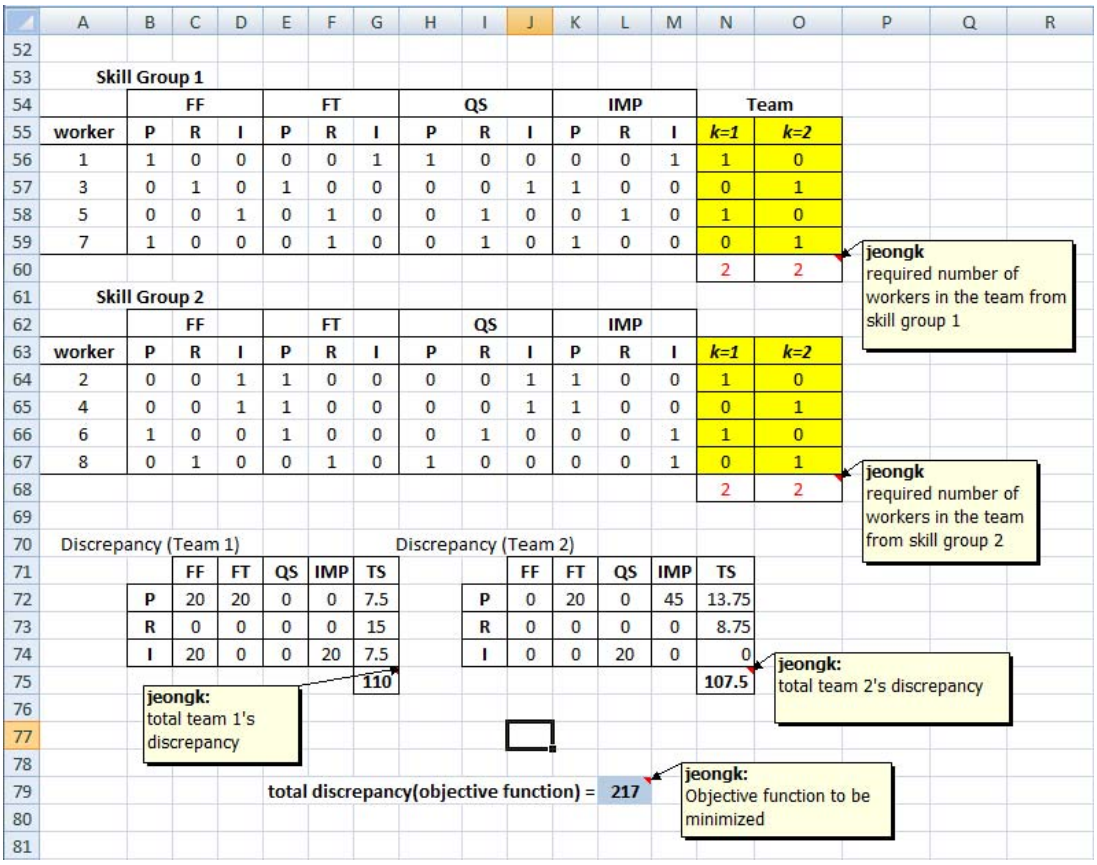
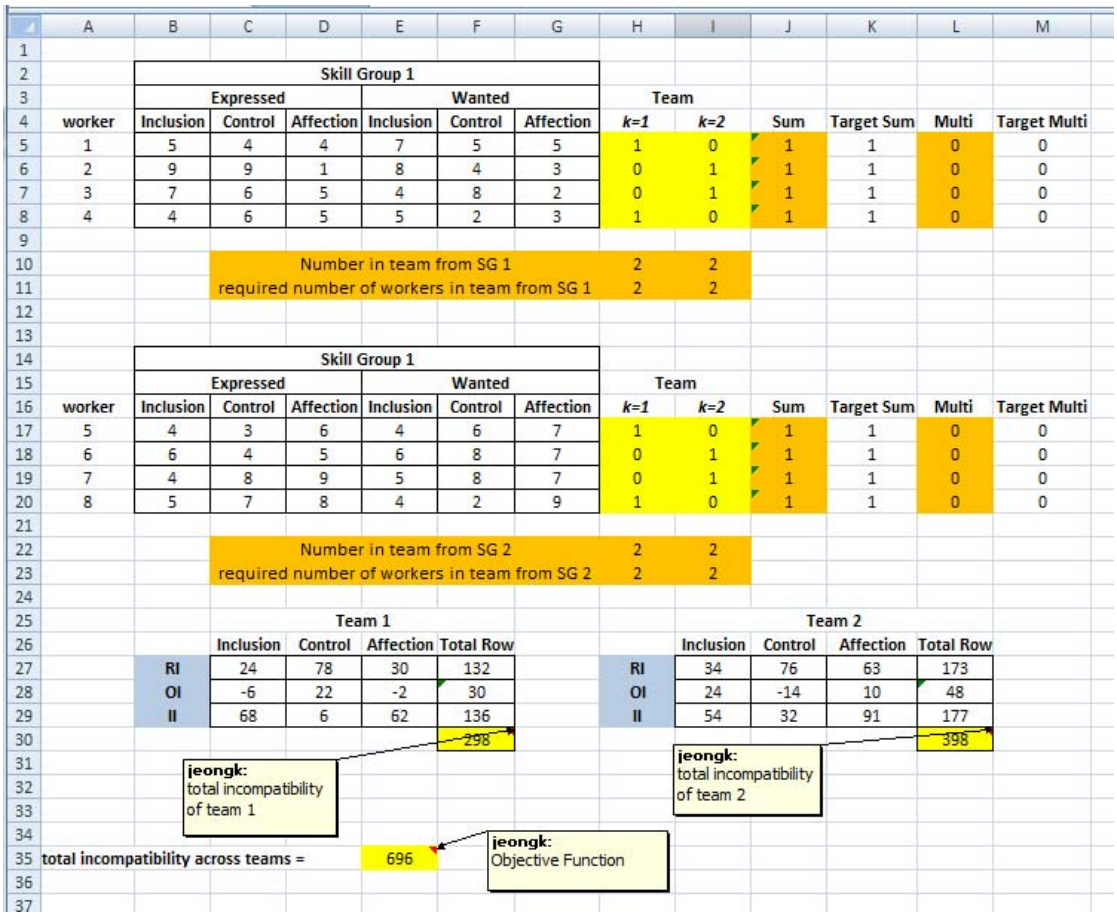


Figure 4. LEPMT Excel Model Screen with Results from Premium Solver

Figure 6 describes the EPS formulation using the spreadsheet model in Figure 5. For example, 'E35'is set as an objective function to be minimized. 'By changing variable Cells:' pane displays the cells for decision variables. As explained in Figure 5, we have two decision variable cells: 'H5:I8' for skill group 1 and 'H17:I20' for skill group 2. The first constraint in 'Subject to the Constraints:' pane is 'H10:I10' = 'H11:I11' to ensure that the number of workers assigned to each team, 'H10:I10', satisfies the required number of skill group 1 workers defined at 'H11:I11', representing that both teams need two workers from the skill group 1. The third constraint, 'H22:I22' = 'H23:I23', describes the same information for skill group 2.

The second and fourth constraints, 'H17:I20' = binary, 'H5:I8' = binary, define that the decision variables are binary for skill groups 1 and 2, respectively. The fifth and seventh constraints, 'J17:J20' <= 'K17:K20' and 'L17:L20' <= 'M17:M20', ensure that one worker from the skill group 1 can be assigned to at most one team by making sure that the sum and multiplication of two corresponding decision variables for each worker should be less than equal to 1 and 0, respectively. The sixth and eighth constraints, 'J5:J8' <= 'K5:K8' and 'L5:L8' <= 'M5:M8' define the same information for workers in the skill group 2.



	A	B	C	D	E	F	G	H	I	J	K	L	M
1													
2		Skill Group 1											
3		Expressed			Wanted			Team					
4	worker	Inclusion	Control	Affection	Inclusion	Control	Affection	*k=1*	*k=2*	Sum	Target Sum	Multi	Target Multi
5	1	5	4	4	7	5	5	1	0	1	1	0	0
6	2	9	9	1	8	4	3	0	1	1	1	0	0
7	3	7	6	5	4	8	2	0	1	1	1	0	0
8	4	4	6	5	5	2	3	1	0	1	1	0	0
9													
10			Number in team from SG 1					2	2				
11			required number of workers in team from SG 1					2	2				
12													
13													
14		Skill Group 1											
15		Expressed			Wanted			Team					
16	worker	Inclusion	Control	Affection	Inclusion	Control	Affection	*k=1*	*k=2*	Sum	Target Sum	Multi	Target Multi
17	5	4	3	6	4	6	7	1	0	1	1	0	0
18	6	6	4	5	6	8	7	0	1	1	1	0	0
19	7	4	8	9	5	8	7	0	1	1	1	0	0
20	8	5	7	8	4	2	9	1	0	1	1	0	0
21													
22			Number in team from SG 2					2	2				
23			required number of workers in team from SG 2					2	2				
24													
25			Team 1						Team 2				
26			Inclusion	Control	Affection	Total Row			Inclusion	Control	Affection	Total Row	
27		RI	24	78	30	132		RI	34	76	63	173	
28		OI	-6	22	-2	30		OI	24	-14	10	48	
29		II	68	6	62	136		II	54	32	91	177	
30						298						398	
31													
32													
33													
34													
35	total incompatibility across teams =				696								
36													
37													

Figure 5. LPEMT-TIC

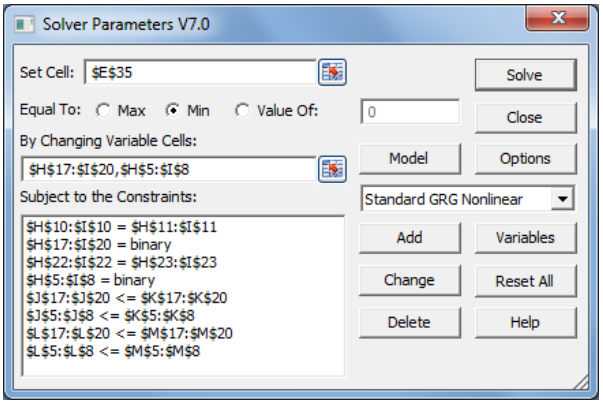


Figure 6. Solver Formulation for LPEMT-TIC

CONCLUSIONS AND FURTHER STUDIES

In this study, we studied three personality models – MBTI, KCI and FIRO-B – in terms of effective team formation. Rather than analyzing the validity of these personality models, we focused on how to use the results from these personality models systematically to create effective teams. Based on quantitative characteristics of outputs, both KCI and FIRO-B seem to be more appropriate than MBTI for this purpose. We also provided the mathematical formulation to use FIRO-B as a decision aid tool to create effective teams. Excel spreadsheet model using Premium Excel Solver was also presented with numerical examples. The diverse validation methods of this mathematical formulation through field studies – for example, the comparison of performance between LPEMT and/or LPEMT-TIC based teams and other teams – will be an interesting future research topic.

REFERENCES

[1] André, M., Baldoquínand, María G. & Acuña, Silvia T. *Formal model for assigning human resources to teams in software projects*. Information and Software Technology, 2011, 53, 259-275.

[2] Fitzpatrick, Erin L. & Askin, Ronald G. *Forming Effective Worker Teams with Multi-Functional Skill Requirements*. Computers & Industrial Engineering, 2005, 48, 593-608.

[3] Furnham, A. & Stringfield, Paul. *Personality and Work Performance: Myers-Briggs Type Indicator Correlates of Managerial Performance in Two Cultures*. Personality and Individual Differences, 1993, 14, 145-153.

[4] Gardner, William L. & Martinko, Mark J. *Using the Myers-Briggs Type Indicator to Study Managers: A Literature Review and Research Agenda*. Journal of Management, 1996, 22, 1996, 45-83.

[5] LaFollette, William R. & Belohlav, James A. *Using Simulation to Study the Effects of Team Compatibility*. Journal of Experiential Learning and Simulation, 1981, 3&4, 181-190.

[6] Myers, I. Briggs. Introduction to type, Palo Alto, CA: Consulting Psychologists Press, 1980.

[7] O'Neill, Thomas A., and Theresa J. B. Kline, *Personality as s Predictor of Teamwork: A Business Simulator Study*. North American Journal of Psychology, 2008, 10, 67-78.

[8] Peeters, Miranda A. G., Van Tuijl, Harrie F. J. M., Rutte, Christel G. & Reymen, I. M. M. J. *Personality and Team Performance: A Meta-Analysis*. European Journal of Personality, 2006, 20, 377-396.

[9] Schutz, W., The interpersonal Underworld: FIRO A Three-Dimensional Theory of Interpersonal Behavior. Palo Alto, California: Science and Behavior Books, Inc., 1966.

CONTROL LIMITS VERSUS ACCEPTANCE LIMITS - WHICH LIMITS ARE APPROPRIATE FOR YOUR TASK?

Donald S. Holmes, Stochos Inc. P.O. Box 247, Duanesburg, N.Y. 12056.
(518) 895-2896, dsholmes@stochos.com
A. Erhan Mergen, Rochester Institute of Technology, Saunders College of Business
Decision Sciences, 107 Lomb Memorial Drive, Rochester, N.Y. 14623-5608.
(585) 475-6143, emergen@saunders.rit.edu

ABSTRACT

The objective of this paper is to discuss and clarify the meaning and use of the different statistical limits used in managing processes so that mistakes that are commonly seen in industrial practice can be avoided. The difference between "control" and "acceptance" will also be discussed. The idea of this paper was triggered by a recent paper written by Henderson [5] in *Quality Progress* journal.

Key words: Statistical limits, process control, process capability, process acceptance.

DISCUSSION

There are several limits in statistical process control (SPC) and they all serve different purposes depending on the question being asked. In order to ensure that proper action(s) can be taken to rectify and/or improve a process, it is critical for the user to do the following: ask the right question, use the right set of limits, and properly determine the limits in order to come up with the right answer. In this paper, we will discuss the basic questions of process control and acceptance:

- stability (in-control),
- performance (capability), and
- acceptance questions

and also clarify the difference between control and acceptance.

Stability (In-control):
The stability question deals with the stability of the process, e.g., are the process average and width stable over time, does the process distribution stay stable over time, etc. In other words, the question being addressed is, *"Are unexpected things happening in my process that I may capitalize on to improve quality?"* The stability question is critical in process management because the future behavior of the stable processes can be predicted with accuracy and, as a result, operations' planning is easier. Keep in mind that the word 'stability' refers to statistical stability, i.e., the only variation displayed by the process is due to random (common) causes.

Control charts are the most effective way of testing the stability question (see, for example, SPC books by Grant and Leavenworth [3], Montgomery [12] for different types of control charts). Limits on the control charts, i.e., *control chart limits*, define the area in which the variation is accepted as due to random (common) causes. As long as sequentially gathered data remains inside these limits, the process has given no indication of instability. Keep in mind that a stable process does not necessarily produce output that is acceptable to the customer(s). This is one of the common misunderstandings in practice.

A stable (in-control) process only means the process is behaving the way we expected it to behave; i.e., there is nothing unusual going on in the process. Such a process may or may not meet the customer expectations, which is a different question to be tested. This issue was also raised by Grant and Leavenworth [3] in their book:

"The word '*control'* has a special technical meaning in the language of statistical quality control. A process is described as 'in control' when a stable system of chance causes seems to be operating. However, the word is often misused and misinterpreted, particularly by those who have been briefly exposed to the jargon of statistical quality control without having had a chance to learn its principles."

Two common misconceptions among practitioners are: if a process is not meeting the customer specifications, it (the process) will generate data for which the control charts indicate that something unusual is happening; or a process which is meeting the customer specifications will generate data for which the control charts indicate nothing unusual is happening!! These two issues, i.e., stability and meeting the customer specifications, must be kept separate.

In addition, one can also test the stability of the process by comparing the two variance estimators without using the control charts. An alternative estimate for the variance is the one computed using mean square successive differences (MSSD). (See, for example, Neumann, et al. [13], Holmes and Mergen ([6, 10].) The MSSD is defined as

$$\text{MSSD} = \frac{1}{(n-1)} \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2 \tag{1}$$

Using these differences an unbiased estimate for the process variance is defined by Hald [4] as

$$q^2 = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2 \tag{2}$$

and the MSSD standard deviation is determined by taking the square root of the q^2.

The variance estimated through MSSD, q^2, as defined above, looks only at the successive differences (by taking into account the time order of the data) and represents the variation that a process could display if some of the non-random elements, such as trends, cycles, etc., were eliminated.

The significance of the difference between the regular and the MSSD variance estimates can be tested using the test given in Dixon and Massey [1]:

$$z = \frac{1 - \frac{q^2}{s^2}}{\sqrt{\frac{n-2}{(n-1)(n+1)}}} \tag{3}$$

where s^2 is the regular variance estimate and n is the number of observations.

z values between ±3 indicate that the difference between the two estimates is not statistically significant, i.e., the process seems to be stable (i.e., in-control). Values (i.e., z values) bigger than +3 and less than -3 indicate that the two variance estimates are significantly different and thus the process is not stable. Values bigger than +3 imply trend and/or long-term cycles in the process and values less than -3 imply short term cycles in the process.

An often neglected issue in determining the control limits on the chart is whether the subgroups are formed *rationally*. Rational subgroups are those which are formed in such a way that the variation within the subgroups is random so that the non-random process changes that take place between subgroups can be detected faster. Otherwise, inflated variation within the subgroups will hide the process changes; and as a result, it will take longer to detect those changes and react to them. This point was emphasized by Shewhart [14], the originator of control charts, as one of the critical factors for the successful use of control charts. However, there may be cases where the formation of the subgroups does not follow the rational subgroup principles outlined above. In some cases, variation within the subgroups may be inflated intentionally given the nature of the process. One example of the reasons for the inflation of the limits is to minimize the unnecessary "out-of-control" signals given by the control chart, if there are natural (and unavoidable) batch-to-batch variation in the process, such as in some chemical processes.

Performance (Capability):
Another critical question tested by SPC is whether we are producing outputs (goods or services) that meet the customer requirements, i.e., the capability question. This question is different from the stability (in-control) question. Based on the answers for the stability and capability questions, a process can be in one of the four states:

- Process in control and capable of meeting the specification limits
- Process in control but not fully capable of meeting the specification limits
- Process not in control but currently meeting the specification limits
- Process not in control and not meeting the specification limits

The proper way of checking whether the process is fully meeting the specification limits is to compare the *natural tolerance limits* of the process against the specification limits set by the customer. Natural tolerance limits define the area where roughly 99.9% of the process output falls. If the process output follows a Normal distribution, the natural tolerance limits would be:

$$\overline{\overline{X}} \mp 3\sigma_X \tag{4}$$

where $\overline{\overline{X}}$ is the estimated process average and σ_x is the process standard deviation estimate. If the process distribution is significantly different from Normal, then the area under the curve which corresponds to roughly 99.9% should be determined by using the proper distribution which describes the process.

A key issue in process capability is to keep the *center* issue and the *width* issue separate. If there is a target value (nominal) given by the customer, the process average needs to be compared to this value to see how close it is to the target, i.e., the center issue. In addition, the width of the process, determined by the natural tolerance limits, can be compared to the customer specifications to see if the process width is narrow enough to meet the specifications. For example, a process may have a very narrow

width to meet the specifications easily; however, if the process average is too far from the target, the process will still generate some unacceptable output.

In practice, quality control practitioners tend to use capability indices, such as Cp, Cpk, Cpl, Cpu, etc., to do the analysis described above. One needs to be careful using these indices since these indices assume that the process follows a Normal distribution. This is another error that is fairly common in practice, i.e., using these indices when the process distribution is not Normal. For capability indices for non-Normal process distribution see, for example, Holmes and Mergen [7].

When the process is not in-control, the analysis described above should be considered to reflect the current performance of the process, i.e., no prediction should be done for the potential capability of the process. In other words, since the process is not stable, we cannot tell what the future process output will be. To determine the potential capability of the process when the process is not in-control, i.e., not stable, one needs to use a capability standard deviation estimate to calculate the capability indices, such as Cp. For example, one estimate for the capability standard deviation is the one determined using MSSD as described above in equation (2). Another capability standard deviation estimate would be the one that takes into account the "runs" that may exist in the process. The run, defined as successive points above or below the median (see Holmes and Mergen [9] for the details), is another variance estimator which takes into account the time order of the process data. This variance estimate represents what the process variance would be if the runs, as defined above, were eliminated. The smaller of the two capability standard deviation estimates, i.e., the estimate through MSSD and the estimate using runs in the process is one way to determine the C_p index which shows the potential capability of the process. The Cp index is defined as:

$$C_p = \frac{USL - LSL}{6\sigma} \tag{5}$$

where USL and LSL are the upper specification limit and lower specification limit, respectively, and σ is the process standard deviation estimated either using the MSSD or the runs in the process.

Acceptance:

The acceptance issue has not received as much attention as the control and capability issues have in SPC. In some instances, processes, because of their nature, are expected to have unavoidable and natural shifts in their average value but are still able to meet the specifications set by the customer. This may occur when the standard deviation of the process, at the various average values of the process, is very small relative to the difference between the upper and lower specification limits given by the customer. Such a process is not considered in-control by SPC standards but may still be able to meet the specification limits. In a situation like this, the proper question to test should not be "Is the process average in-control?" but rather "Is the process producing output satisfying the customer specifications?" Checking the stability of the process average for such a process would not generate any valuable information for process managers since the answer would be obvious, i.e., "not in-control." In other words, using control chart limits is not the right thing to do in this case. What we want is protection against the case where the process average may deviate so much from its desirable value that it will start producing some non-conforming output, i.e., output that does not lie within customer specification limits.

We could check this through the use of acceptance sampling plans, such as MLT 105 and 414. However, this would be too late. We would like to catch defective output in the process before the quality system audit at the end. In other words, we should move the acceptance procedure on-line with charting similar to control charts, i.e., acceptance charting. Acceptance charting is similar to control charting but acceptance charts address a different question: *Is my process producing defective (nonconforming) output?* The purpose of the acceptance charts is to evaluate a process in terms of whether or not it could be expected to satisfy output specifications. Again, the question that is being tested under acceptance charts is *"Is the process producing acceptable parts?"* <u>not</u> *"Is the process in statistical control?"* We will discuss how to set the maximum/minimum acceptable values for the process average along with the upper and lower acceptance limits for the sample averages using acceptable quality level (AQL) as the design criteria. For more detail on acceptance charts see, for example, Freund [2], Montgomery [12], Holmes and Mergen [8, 11]. The steps for building acceptance chart are as follows:

1. First, decide on the AQL that you think appropriate for the process.
2. Determine a maximum allowable value for the process mean, known as the Upper Acceptable Process Mean (UAPM), at $k_1\sigma_x$ below the upper specification limit (USL), where σ_x is the capability standard deviation of the x's. The capability standard deviation is the standard deviation estimate which is independent of changes in the process average, such as $\bar{R}/d_2$ or $\bar{s}/c_4$. $\bar{R}$ and $\bar{s}$ are the average of the subgroup (sample) ranges and the average of the subgroup standard deviations, respectively. Factors d_2 and c_4 are the correction factors for a given subgroup (sample) size. (The MSSD standard deviation could be another estimate for the capability standard deviation.) For example, if the x values are roughly normally distributed, then using k_1=3 will produce an AQL value of approximately 0.15% for a one-sided specification limit.

$$UAPM = USL - k_1\sigma_x \tag{6}$$

3. Add $k_2\sigma_{\bar{x}}$ to UAPM to arrive at the Upper Acceptance Limit (UAL) for the sample averages, where $\sigma_{\bar{x}}$ is the standard deviation of $\bar{x}$'s, i.e., sample averages. The value of k_2 sets the probability of the acceptance of output that has a quality level of AQL. For example, using k_2 = 3 would set the probability of accepting output that has a quality level of AQL to approximately 0.9985 for a one-sided specification limit.

$$UAL = MAPM + k_2\sigma_{\bar{x}} \tag{7}$$

The acceptance limit for the lower specification limit (LSL) case is done same way as follows.

$$LAPM = LSL + k_1\sigma_x \tag{8}$$

and

$$LAL = LAPM - k_2\sigma_{\bar{x}} \tag{9}$$

where LAPM is the lowest acceptable process mean and the LAL is the lower acceptance limit for

the sample $\bar{x}$.

Note that the standard deviation of the $\bar{x}$'s is the standard deviation of the x's divided by the square root of the sample size, i.e., $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$ where n is the sample size. Thus an $\bar{x}$ value exceeding the UAL or LAL would indicate that the process average may have shifted above/below the level that is tolerated so the process may produce unacceptable output. When the $\bar{x}$ values (i.e., sample averages) stay within the established UAL and LAL, the process would be considered (i.e., accepted) to be producing acceptable output. Note that the acceptance chart can also be used for one-sided specifications; in that case there would be one acceptance limit, either UAL or LAL.

CONCLUSION

In conclusion, in this paper we discussed several key issues in SPC, along with the proper statistical limits to check those issues. It is critical that not only the right kind of limits be used but also that those limits be determined properly given the SPC question that we are trying to answer.

REFERENCES

[1] Dixon, W.J. and Massey, F.J. Introduction to Statistical Analysis. New York, NY: McGraw-Hill, 1969, 353-354.

[2] Freund, R.A. *Acceptance Control Charts*. Industrial Quality Control, 1957, 14(4), October, 13-23.

[3] Grant, E.L. and Leavenworth, R.S. Statistical Quality Control, sixth edition, New York, NY: McGraw-Hill, 1988.

[4] Hald, A. Statistical Theory with Engineering Applications. New York, NY: Wiley, 1952, pp. 357-360.

[5] Henderson, R.K. *Know Your Limits: Avoid Misunderstandings and Mismanagement of Statistical Limits in Manufacturing Settings.* Quality Progress, 2011, 44(5), May, 36-41.

[6] Holmes, D.S. and Mergen, A.E. *An Alternative Method to test for Randomness of a Process*. Quality and Reliability Engineering, International, 1995, 11(3), 171-174.

[7] Holmes, D.S. and Mergen, A.E. *Problems with Non-Normal Distributions and Process Performance Measures*. The 1997 Northeast Decision Sciences Institute Meeting Proceedings, 1997, Annapolis, Maryland, April 2-4, 299-301.

[8] Holmes, D.S. and Mergen, A.E. (2000). *EWMA Acceptance Charts*. Quality and Reliability Engineering, International, 2000, 16(2), 1-4.

[9] Holmes, D.S. and Mergen, A.E. *Estimating Potential capability of an Unstable Process*. The 2004 Northeast Decision Sciences Institute Meeting Proceedings, 2004, Atlantic City, NJ, March 24-26, 290-292.

[10] Holmes, D.S. and Mergen, A.E. *Using Descriptive Statistics in Statistical Process Control*. The 2007 Northeast Decision Sciences Institute Meeting Proceedings, 2007, Baltimore, Maryland, March 28 – 30, 566-571.

[11] Holmes, D.S. and Mergen, A.E. *Use of Acceptance Charts in the Six Sigma Environment*. The 2009 Northeast Decision Sciences Institute Meeting Proceedings, 2009, Uncasville, CT, April 1-3, 443-448.

[12] Montgomery, D.C. Introduction to Statistical Quality Control. (fourth edition) New York, NY: Wiley, 2001.

[13] Neumann, J.V., Kent, R.H., Bellinson, H.R. and Hart, B.I. *The Mean Square Successive Difference*. Annals of Mathematical Statistics, 1941, 12, 153-162.

[14] Shewhart, W.A. Economic Control of Quality of Manufactured Product. New York, NY: D. Van Nostrond, 1931.

IMPROVING PROJECT MANAGEMENT DECISION MAKING BY MODELING QUALITY, TIME, AND COST CONTINUOUSLY

Matthew J. Liberatore, Villanova University, 610-519-4390, matthew.liberatore@villanova.edu
Bruce Pollack-Johnson, Villanova University, 610-519-6926,
bruce.pollack-johnson@villanova.edu

ABSTRACT

The value of a project to a client can be measured in part by the level of quality associated with the completed project. Quality is acknowledged to be an important component of project management, but its *joint* relationship with time and cost previously has not been modeled. This paper introduces the notion of a quality function for individual tasks and uses a bivariate normal function to model quality at the task level. Using two case studies, a translation agency and a software development company, the quality function is specified and incorporated into a mathematical programming model that allows quality to be explicitly considered in project planning and scheduling and addresses the nonlinear tradeoffs between quality, time, and cost. An alternative model formulation leads to the creation of quality level curves that enable managers to evaluate the tradeoffs between quality, time, and cost across each of the example projects. The results of these analyses lead to specific decisions about the planned values for these three fundamental dimensions at the task level and provide insights for project planning and scheduling that can be gained through improved understanding of the choices and tradeoffs.

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 then 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 managing the relationship between time and cost, with little if any attention directed to quality. However, in many situations there are alternative approaches for completing each task, each having its own time, cost, and quality considerations. Differences in quality can arise due to bids offered by competing subcontractors to complete specific tasks. Even different bids by the same subcontractor could imply different quality levels. At the same time, subcontractors might have some flexibility with time and cost to allow for different quality levels of the same task. This can also be true for alternative work plans offered in-house. For example, in completing a foundation for a building there are choices related to the depth of the excavation and the compressive strength of the concrete used. Each of the possible alternatives will achieve different levels of time, cost, and quality associated with this task.

In this paper we introduce the notion of a continuous quality function for individual tasks, specified in terms of time and cost. We provide two case studies, document translation and software development, to demonstrate how the quality function can be specified. We evaluate the application of the bivariate normal function to represent the quality-time-cost relationship. We then incorporate quality at the task level into an analytic model for project planning and scheduling. Using our project examples, we show how the problem of maximizing the quality of the weakest link subject to time and cost limits can be formulated and solved. An alternative formulation leads to the creation of iso-quality curves for these examples that enables managers to evaluate the tradeoffs between time, cost, and quality across each project. Using either formulation, the results lead to specific decisions about the planned values for these three dimensions for each project task. We also offer some insights for project planning and scheduling that are derived from the analysis through improved understanding of these choices and tradeoffs.

LITERATURE REVIEW

In the field of project management, "Quality management has equal priority with cost and schedule management" [1]. This statement makes inherent sense, since project management is concerned with not only managing cost and schedule, but also the actual work completed in order to achieve the project goal. The quality of the work completed then is an important project outcome, since it directly relates to the value of the project deliverables. The Project Management Body of Knowledge [2] has adopted the ISO 9000, clause 3.11 definition of quality as "the degree to which a set of inherent characteristics fulfills requirements" [3]. Quality issues must be addressed in both the management of the project and the product of the project [2]. Specifically, the PMBOK suggests that quality must be addressed throughout the project life cycle, beginning in the project planning phase and continuing through quality assurance and quality control [2]. Unfortunately, no guidance is provided in terms of how quality can be measured in a project context.

Pollack-Johnson and Liberatore [4] illustrate how the quality of a task option can be determined using the Analytic Hierarchy Process or AHP [5]. They extend the discrete time – cost problem by developing a mathematical programming model that determines optimal discrete options defined in terms of time, cost, and quality combinations for specific tasks to maximize overall project quality subject to time and cost constraints.

An alternative approach for measuring quality at the task level is to make a direct subjective assessment of quality [6], [7]. The approach proposed in [6] extends the standard time – cost tradeoff analysis by assuming that quality depends only on time and is independent of cost for a given time, a major limitation of their approach. Project quality is measured as the arithmetic or geometric mean of the quality of the activities or as the minimum quality of the activities. Iso-curves are used to show the relationships between two of the three factors, with the other held at constant levels. Babu and Suresh [6] state: "This paper illustrates the concept in a general project setting using linearity assumptions. More complex relationships could be modeled in a similar fashion to capture reality more closely."

Khang and Myint [7] applied the method provided in [6] to an actual cement factory construction project. In both papers the quality levels associated with all normal times was set at 1.0, while the lower quality level at the crash time for each activity was measured relative to this level. The linearity assumption between quality and time was found to be problematic. The most difficult and controversial task was to assess the quality reduction associated with crashing [7]. Kang and Myint [7] suggested the need for a more holistic measurement of performance quality and a more realistic model to describe the relationships between the quality of individual activities (and therefore of the whole project) and the budget and time allowed.

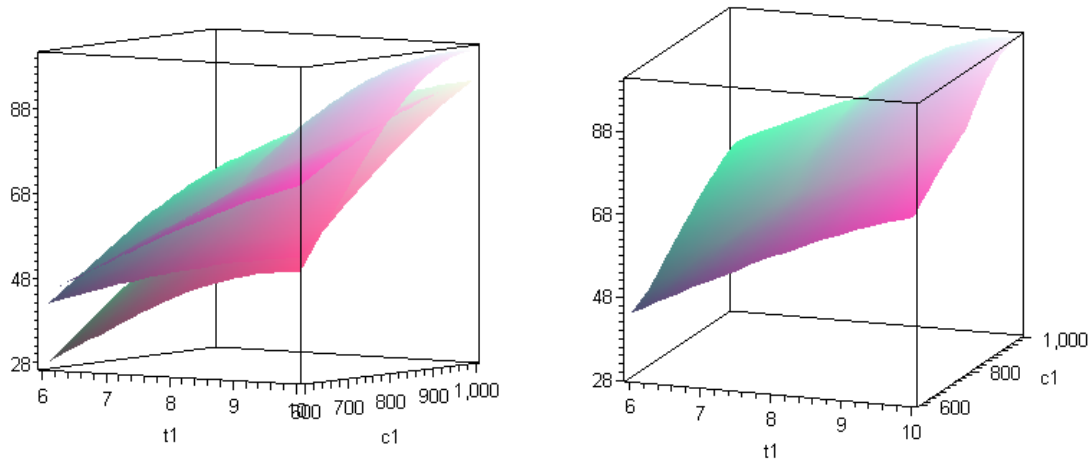
In response to the suggestion in [7] and the limitations of the approach in [6] noted above, in this research we model quality at the task level as a ***continuous*** nonlinear function of *both* cost and time.

THE QUALITY FUNCTION

Development

Our overall goal is to find a rich method for modeling the relationship between quality, time, and cost at the individual task level within a project, as well as at the overall project level. Our approach begins by formulating a model of the quality of each task as a function of the time (t) and cost (c) allocated to it. We assume that there could be different entities who could do the task (individuals or groups, subcontractors, etc.), 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. This approach is consistent with the assumption of efficiency in new product development [8]. Figure 1 shows an example for three different entities that could perform a particular task. Graph (a) shows three individual quality surfaces for a given task, and graph (b) shows the upper envelope of these three surfaces, which we consider the overall quality surface for this task.

Figure 1: Quality graphs



(a) 3 individual quality graphs

(b) the upper envelope overall quality graph

We assume that this overall quality function for a task has two basic properties:

- Holding time constant, quality is an *increasing* function of cost. Thus, if time is fixed, we assume that spending more money on the task will increase quality.
- Holding cost constant, quality is an *increasing* function of time. Thus, if cost is fixed, we assume that spending more time on the task will increase quality.

If we normalize quality to be on a 0-100 scale, and limit time and cost to reasonable (possibly fairly extreme) values for the task at hand, based on the two non-decreasing assumptions above, we would expect the graph to show 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 time/cost tradeoff curve that is decreasing and convex. Thus, to maintain the same level of quality, to reduce the time, one has to pay increasingly more money per unit, such as in standard project activity *crashing* [9]. 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 (such as the individual quality curves in Figure 1a).

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. The probability function is normalized by the area under the surface being 1; our quality function is normalized so that the maximum time (μ_t) and cost (μ_c) values considered reasonable correspond to a quality of 100, or some other constant indicating the maximum quality level possible, *K*. 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

$$Q(t,c) = Ke^{-\left[\left(\frac{t-\mu_t}{\sigma_t}\right)^2+\left(\frac{c-\mu_c}{\sigma_c}\right)^2\right]} \qquad (1)$$

The quality curves in Figure 1 are of this form. For simplicity, we have normalized the initial constant to *K* and eliminated the ½ in the exponent (which means that each σ in the quality function definition would be divided by $\sqrt{2}$ to be interpreted as the usual σ in the bivariate normal distribution). Notice that 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). Notice also that the upper envelope graph may not be smooth (e.g., in Figure 1b, there are some "seams" where the individual surfaces intersected at an angle to each other, at which points the function is not differentiable in certain directions), but we are assuming that we can fit a smooth function that is a reasonable estimate of the upper envelope. For a given value of quality using this bivariate normal functional form, there is a nonlinear relationship between time and cost, generalizing the standard linear approach. This quality function is *increasing* with respect to both time and cost (holding the other constant), consistent with the assumption of efficiency in new product development [8]. Notice that the version of the bivariate normal distribution upon which this functional form is based 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/interaction between the variables.

.

Estimation

In situations where n bids, alternative work plans, or scenarios 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:

$$\text{Minimize } \sum_{j=1}^{n} \{Q(t_j, c_j \mid \mu_{t_j}, \mu_{c_j}, \sigma_{t_j}, \sigma_{c_j}) - q_j)^2 \tag{2}$$

It is necessary to add the constraints:

$$\begin{aligned} \mu_{t_j} &\geq t_j, \ j = 1, 2, ..., n \\ \mu_{c_j} &\geq c_j, \ j = 1, 2, ..., n \end{aligned} \tag{3}$$

because the μ parameters are upper bounds on the values of the t and c variables.

Data Fitting

Translation Example

A translation agency has received a large translation project that needs to be completed in at most 4 days. The project is too large for one translator to complete in this amount of time, so the project manager plans to divide the job between two translators. The project manager has two translators in mind, and has defined a quality scale that connects naturally to two different systems of measuring the quality of translation work. The Interagency Language Roundtable (ILR) provides one of these systems, and has four levels (4, 3, 2+, and 1+) [10]. Another system is the rubric used by graders of the American Translators Association (ATA) certification exams, which identifies "points" for weaknesses or errors in a translation, and categorizes a translation into four categories based on the number of points identified (from 0 to 46+) [11]. The hybrid quality scale system developed was on a 0-44 scale (44 being the highest possible quality), so that (roughly) quality levels that rounded off to 40 (to the nearest 10) would correspond to ILR Level 4, rounded to 30 would be Level 3, rounded to 20 would be Level 2+, and below 15 would be Level 1+. For the ATA rubric, the number of points can be subtracted from 44 to get a rough equivalent value on the 0-44 quality scale. After the subtraction 35-44 points corresponded to Strong in that system, 27-34 Acceptable, 14-26 Deficient, and 13 or less Minimal. Notice that the categories correspond quite closely between the two systems, which help to validate this approach.

The translation agency tended to establish a rate per word with each translator, but for rush jobs, the word rate could be increased, or a flat dollar bonus amount could be added, or a combination of these. Thus the cost of each of the two tasks is divisible (continuous). Similarly, the deadline (time) is divisible. To reflect the reality of the translation world, time is measured in days, thinking of 1 day as 16 working hours (say, 8am to midnight). If the project manager was negotiating on a Monday afternoon at 4pm, then a deadline of 3 days would mean Thursday at 4pm obviously, but 3.5 would correspond to Friday 8am.

With these conventions and the specific translators in mind, the project manager estimated the quality of translation that could be expected for different combinations of deadline and cost for each of the two pieces of the translation (each piece being assigned to a specific translator). The results are shown in Table 1. Notice that Translator 1 has consistently higher quality than Translator 2; this will become significant later.

Table 1: Time, Cost, and Quality Estimates for Translation Project

Translator 1			**Translator 2**		
Deadline (days)	Cost ($)	Quality (0-44)	Deadline (days)	Cost ($)	Quality (0-44)
t	c	q	t	c	q
4	1435	41	4	653	32
3	1435	35	3.75	653	26
3	1750	41	3.75	800	32
3	2150	43	3.75	975	33
2.5	2150	37	3.5	975	26
2	2150	33			

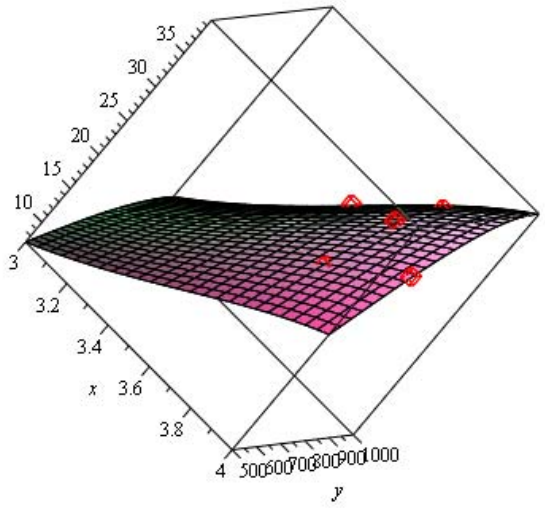
The four parameters of the bivariate normal quality functions for both translators were determined by applying equations (2) and (3) to the data given in Table 1 using Lingo's Global Solver [12]. For Translator 2 the constant K in equation (1) was replaced by 44 to fit the specific translation quality scale developed above. The results are shown in Table 2. The SSE is very small (1.89) and all estimated quality values are within one point of the actual values, well within the measurement error of the scale itself. Figure 2 provides a plot of Translator 2's quality function and actual values from Table 2.

Table 2: Results of Translator Bivariate Normal Quality Function Estimation

	Translator 1	Translator 2
Constant [equation (2)]	50	44
SSE	1.89	1.26
Time mean (μ_t)	4.5086	4.3768
Time std. dev. (σ_t)	3.8223	1.2013
Cost mean (μ_c)	2150.0	975.00
Cost std. dev. (σ_c)	1655.1	671.10

A similar procedure was used for Translator 1, but in using a constant of 44, the best SSE was 11.65. By using a constant value of 50, the SSE was reduced to 1.26, again with no deviations more than 1. The estimated parameters are given in Table 2, and the plot of Translator 1's quality function looks very similar to that for Translator 2. In practice the quality function for Translator 1 could be capped (truncated) at 44 since this is the maximum value on the translator quality scale, if necessary.

Figure 2: Plot of Translator 2's Bivariate Normal Quality Function with Data



Software Development Example

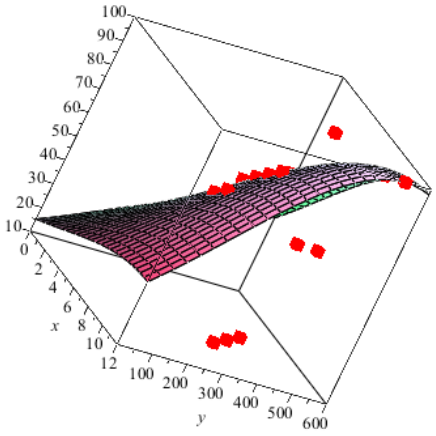
As a second example of fitting a quality curve, we present a case study of a firm that develops software applications for clients. The process typically involves gathering the customer requirements (after making sure the rough, likely cost is in the feasible range for the client), then developing a quote for the project cost. If the quote is accepted, staff are allocated, there is an initial team meeting, the main database and software configuration and testing are pursued in parallel (as well some tasks that are outsourced), then these pieces are integrated and finalized, and there is a final task of analyst testing before the software is sent to the client for their testing.

In working with the manager responsible for these projects, we decided to base the quality level on the concept of expected rework cycles resulting from client testing. Rework based on the client changing their requirements is not considered a quality issue, but rework needed to meet those requirements is. From the point of view of the software development company, clearly rework cycles based on customer testing are damaging to their reputation and future thriving as a business. The manager indicated that their goal as a business is to achieve a high level of reliability, the major component of which is avoiding client rework cycles. To define a 0-100 scale, then, the manager estimated the expected damage from rework cycles, with a client rework cycle costing 20 points on the 0-100 scale for each substantial issue discovered. For example, if the time and cost allocated to a task were estimated to lead to an 80% chance of a client rework cycle for a single major issue, the estimated quality would be $100 - (0.8)*20 = 84$.

The process used to fit the quality curve was similar to the translation example above, but was more complicated because for some tasks there were more than one staff person who could perform the task, with different time, cost, and quality characteristics. This required application of the upper envelope concept discussed earlier. If some time, cost, quality combinations were too inefficient, they were eliminated, and the curve was fit to the upper envelope points only. Figure 3 illustrates this for one of the tasks. The data points under the curve were removed when

fitting the model, since the employee they represent would not be chosen for that task (the quality is too much lower than the other ways of performing the task for the same time and effort). The results of the fits for the seven tasks are shown in Table 3.

Figure 3: Fit for Task 4 (Standard Configuration and Testing) Showing Upper Envelope



Notice that tasks 2 and 3 (and 0 and 8, the dummy start and finish activities) have fixed time and cost, so no decision needs to be made about them.

Comparison with the Babu and Suresh approach

The approach of Babu and Suresh [6] and Khang and Myint [7] assumes quality is only a linear function of time. However, cost is also assumed to be a linear function of time, so that if we plot quality as a function of both time and cost, the 3-D graph would just be a line segment. As an example, Figure 4 shows the 3-D graph of quality vs. time and cost for Translator 2 using the estimated data from Table 1, as well as the two endpoints of the line segment that is the implied quality function in [6], assuming a normal (time, cost quality) of (4,653,32) and crash values of (3.5,975,26).

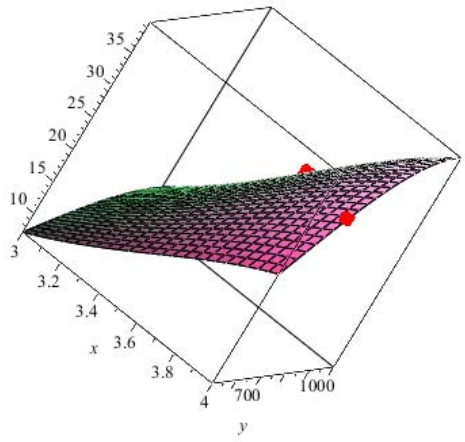
Clearly, just assuming that Translator 2's quality can only lie on the line segment between the two highlighted points is much more restrictive than having the full continuum of possibilities of the curved surface, and does not reflect the full range of choices.

Considering the software example, it is not clear how the method in [6] could be implemented, since the nature of the tradeoffs did not involve crashing. Instead, the empirical data indicated that later deadlines for the same number of billable hours yield higher quality. This relationship is captured by the bivariate quality function. However, the assumptions of [6] would have resulted in a line segment that did not even line up with the quality surface, and so the results would be misleading.

Table 3: Bivariate Normal Quality Function Parameters for Software Project Tasks

Task No.	Description	Pred.	Time Mean	Time Std. Dev.	Cost Mean	Cost Std. Dev.
0	Dummy	--	0	0	0	0
1	Requirements & Review	0	4.33	10.42	204.00	116.24
2	Setup	1	4.00	0	1045.00	0
3	Vendor Configuration & Testing	2	1.00	0	1000.00	0
4	Standard Configuration & Testing	2	12.00	12.82	525.00	520.65
5	Data Base Configuration & Testing	2	12.55	11.16	5493.10	4004.92
6	Development, Integration & Completion	3, 4, 5	1.54	1.88	56.16	52.06
7	Quality Assurance Testing	6	1.00	1.28	381.64	1246.51
8	Dummy	7	0	0	0	0

Figure 4: Comparison of Fitted Bivariate Normal Quality Function and Babu and Suresh Quality Function for Translator 2



MODEL FORMULATION

Maximizing Minimum Quality

Using the estimated quality functions for each project activity, we can determine the tradeoffs between quality, time, and cost. 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 successor of any activity, and that the finish activity (activity N+1 , also a dummy activity) is the only activity that has no successors. It is common to use predecessors rather than successors for project scheduling formulations of this type, but for this example, the formulation turns out to be much more concise and elegant using successors.

Define the following parameters and variables:

t_i = the duration of activity i, for $i = 1,\ldots,N$
c_i = the cost of activity i, for $i = 1,\ldots,N$
q_i = the quality of activity i, for $i = 1,\ldots,N$
S_i = the set of activities that are immediate successors of activity i, for $i = 0,\ldots,N$
T_{UB} = upper bound on the total project time
C_{UB} = upper bound on the total project cost
s_i = the scheduled start time for activity i, for $i = 0,\ldots,N+1$
$t_{\min_i}$ = lower bound on the duration of activity i, for $i = 1,\ldots,N$
$c_{\min_i}$ = lower bound on the cost of activity i, for $i = 1,\ldots,N$

Relevant quality measures from a project perspective 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:

$$Q_{\min} = \min_{1 \le i \le N} q_i \tag{4}$$

The quality-time-cost problem can be thought of as having three objectives, and can be modeled as a goal programming problem. In our formulation, we first maximize $Q_{\min}$ while setting an upper bound on total project cost and total project time. The nonlinear program is given as equations (5) – (17) below:

$$\text{Maximize} \quad Q_{\min} \tag{5}$$

subject to:

$$Q_{\min} \le q_i, i = 1,2,,...N \tag{6}$$

$$q_i = Q_i(t_i, c_i) = K * \exp\left\{-[(t_i - \mu_{t_i})/\sigma_{t_i})]^2 - [(c_i - \mu_{c_i})/\sigma_{c_i})]^2\right\}, i = 1,2,,...N \tag{7}$$

$$\sum_{i=1}^{N} c_i \le C_{UB} \tag{8}$$

$$s_0 = 0 \tag{9}$$

$$s_k \ge s_i + t_i \qquad \forall i = 0,...,N, \forall k \in S_i \tag{10}$$

$$s_{N+1} \le T_{UB} \tag{11}$$

$$s_i \ge 0 \qquad \forall i = 1,...,N+1 \tag{12}$$

$$t_i \ge t_{\min_i}, i = 1,2,...,N \tag{13}$$

$$c_i \ge c_{\min_i}, i = 1,2,...,N \tag{14}$$

$$t_i \le \mu_{t_i}, i = 1,2,...,N \tag{15}$$

$$c_i \le \mu_{c_i}, i = 1,2,...,N \tag{16}$$

$$q_i, t_i, c_i \ge 0, i = 1,2,...,N \tag{17}$$

Equation (6) is a modeling device to obtain the minimum of the q_i . The mathematical program is nonlinear due to Equation (7), the bivariate normal quality function. Equations (8) and (11) set bounds on project cost and completion time, respectively. Equation (10) is needed to enforce the successor relationships among the tasks in the network. Lower bounds on the time and cost variable values are set by equations (13) and (14). These bounds are usually based on the smallest completion time and cost possible for each activity. Equations (15) and (16) reflect the upper bounds on the cost and time for each activity imposed by the definition of the quality function, although other values for the upper bounds could be considered. The independent decision variables are the time and cost for each project activity (which determine the quality for each activity). This problem can be solved using Lingo's global solver [12] and generalizes the standard time/cost tradeoff problem [9].

The Minimum Cost Formulation

An alternative formulation of the problem considers minimizing total project costs with bounds on project completion time and quality. In the problem formulation (5) – (17), the objective in (5) becomes minimizing total cost or $\sum_{i=1}^{N} c_i$, thus eliminating equation (8), and then specifying a lower bound for project quality in equation (6), which becomes

$$Q_{LB} \le q_i, i = 1, 2, ,...N \qquad (18)$$

Analyzing the minimum cost model for different total project times T_{UB}, we can find the minimum cost possible that finishes the project within a given time and maintains a minimum quality of at least the lower bound. Note that this is closely related to the traditional crashing problem, because both work with a fixed level of quality. A set of iso-quality (level) curves can then be constructed corresponding to different fixed quality levels, as will be demonstrated for our two examples below. These iso-quality curves assist managers in evaluating the tradeoffs between time, cost, and quality across the project. This model then generalizes the standard time-cost problem by including quality considerations.

RESULTS

Translation Project Example

Maximizing Q_{min}

Returning to the translation example, consider the situation requiring a decision at 4pm Monday concerning how much to offer to pay each translator, and what deadline to give each of them, to maximize the overall quality of the job (as defined by $Q_{\min}$) and have the results in hand by noon Thursday (equivalent to 3.75 days later), at a cost of no more than $2400. This is a special case of the quality-time-cost model (equations 5 – 17) with two parallel activities. Using the bivariate normal quality function estimates the model was solved using Lingo's Global Solver [12] with the results given in Table 4. We note in advance that all computational times for all runs using both the max $Q_{\min}$ and min cost models for both case studies was less than 0.05 seconds.

It was noted before that Translator 1 was of higher quality than Translator 2. Since the objective is to maximize $Q_{\min}$, the agency should put as much money and time into Translator 2 as possible to bring the quality up as high as possible. The lower bound on the cost for Translator 1 is \$1435, which means the upper bound for Translator 2 is \$2400-\$1435=\$965, so both Translator 2 variables are at their upper bounds (\$965 and 3.75 days). The cost allotted to Translator 1 is the minimum \$1435, and the time is set at the value that will achieve the exact same quality as Translator 2 (33.51 on the 0-44 scale), which turns out to be 2.74 days (Wednesday noon). It is interesting to note that it is very possible that there is no real advantage to the agency to set a tighter deadline for Translator 1 than the 3.75 days being given to Translator 2. In fact, the agency could make the deadline for Translator 1 to also be 3.75 days, without changing $Q_{\min}$ (this is an alternate optimal solution, clearly preferable to the agency!).

Table 4: Solution of Translator Project Example Using Bivariate Normal Quality Function

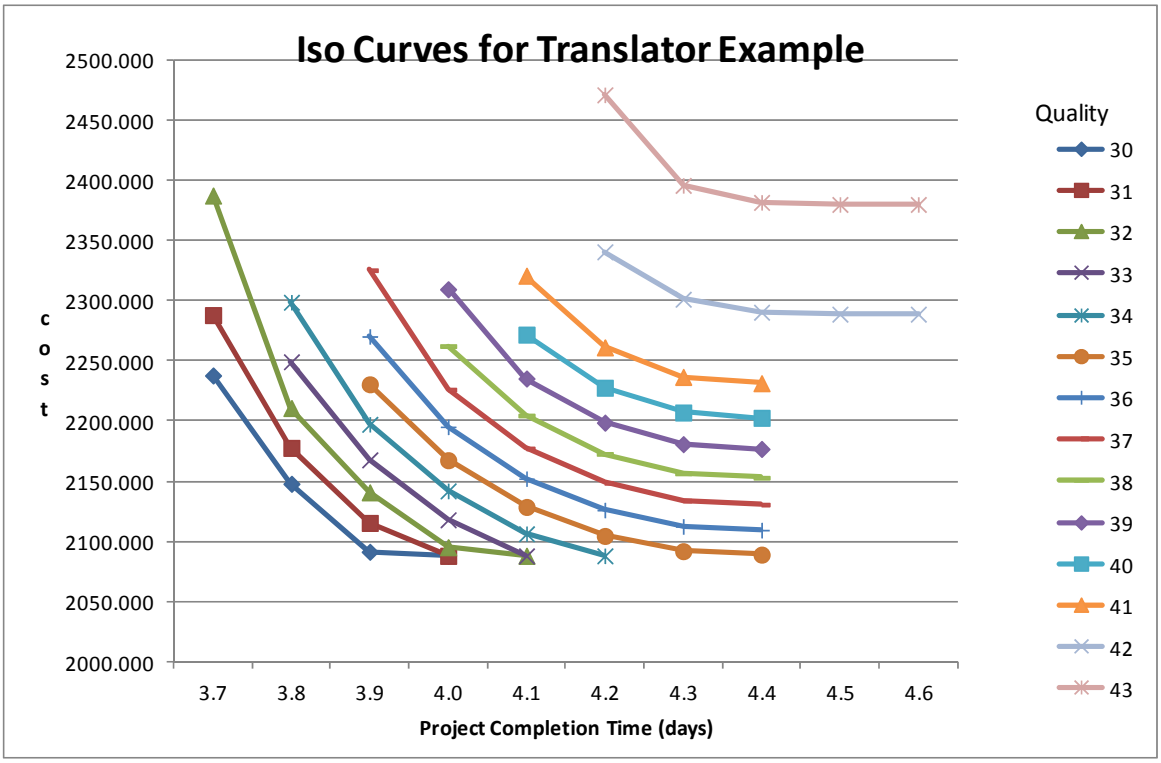
	Translator 1	Translator 2	Project
Time	2.74	3.75	3.75
Cost	1435.00	965.00	2400.00
Quality	33.51	33.51	33.51

Iso-Curves

Using the minimum cost formulation discussed above, iso-curves were constructed for the translator project and are shown as Figure 5. Notice that the iso-curve for a higher quality level lies above and to the right of that for a lower quality level. Each iso-curve in Figure 5 provides the concave time-cost tradeoffs traditionally discussed in the project management literature. The iso-curves in Figure 5 provide a concise summary of the relationship among quality, time, and cost, and can be used by the project manager to make well-informed decisions about how to execute the project in terms of specific decisions about the work plans for all project tasks.

There are several interesting observations that can be gleaned from Figure 5. First, assuming that the project must be completed by a deadline of 4 days, then a quality of 32 can be achieved at about the minimum cost of just under \$2100. Starting from this solution, total cost will increase by the relatively small amount of around \$50 for each unit increase in quality, and would seem to be worthwhile. However, the maximum achievable quality at the deadline is 39, regardless of cost. Therefore, to achieve higher levels of quality the deadline must be extended. Starting from a given quality and cost at the deadline, relaxing the deadline just a bit (.1 day) can lead to an increase in quality by a point and no change or a reduction in cost as well. Second, to achieve the very high quality levels of 42 or 43 the project deadline must be extended and costs will increase significantly as well, indicating that these are not likely options. Third, the fastest time for project completion is 3.7 days but then quality can be at most 32 for a relatively high cost \$2400. However, if time is increased by a small increment to 3.8 days, quality can increase to 34 with a *reduction* of cost to \$2300, indicating extreme sensitivity to time on the low end.

Figure 5: Iso-Quality Curves for Translator Example



Software Development Example

Minimizing Cost

Table 5 and Figure 6 shows the minimum cost problem solution for the software example, using $T_{UB} = 19$, $Q_{LB} = 69$, and the minimum values for the time and cost of each task set at 30% of the mean value (maximum). Recall that tasks 2 and 3 have fixed time and cost values, so have values that are whole numbers (and quality values that were evaluated to be 100).

Notice that most of the quality values are 69, which is the minimum quality level, Q_{LB} . Tasks 2 and 3 have quality of 100, because their values were fixed. However, task 7 has a quality of 70.7535, which corresponds to the quality when the time and cost values are set at their minimum (notice that the values are 30% of the mean values for both). This is a special case of a general result that can be easily shown:

Theorem: For the minimum total cost formulation described above, if the problem is feasible, then there will exist an optimal solution in which

$$q_i^* = \max\left\{Q_{LB}, Q_i\left(t_{\min_i}, c_{\min_i}\right)\right\}, \forall i$$

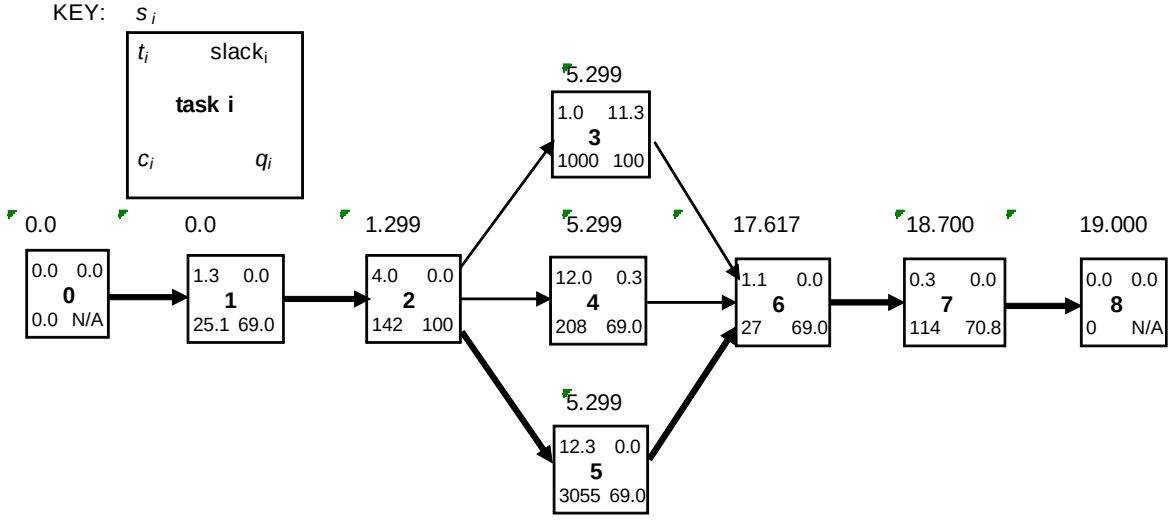
Intuitively, this theorem simply says that the required quality level will be achieved exactly if possible, but if the minimum time and cost force a higher quality, the time and cost will be set at those minimum values (notice that this applies to tasks 2 and 3 as well). If a task has slack at a given optimal solution that fits the theorem, its time value might be able to be increased, without

increasing the cost, which would yield a higher quality for that task as an alternative optimal solution (and clearly preferable), similar to what we discussed for the maximizing Q_{min} model above.

Table 5: Software Project Minimum Cost Solution for $T_{UB} = 19$, $Q_{LB} = 69$

Task No.	Description	Time	Cost	Quality
1	Requirements & Review	1.29897	141.780	69.0000
2	Setup	4.00000	1045.00	100.000
3	Vendor Configuration & Testing	1.00000	1000.0	100.000
4	Standard Configuration & Testing	12.0000	207.845	69.0000
5	Data Base Configuration & Testing	12.3180	3054.89	69.0000
6	Development, Integration & Completion	1.08295	27.1234	69.0000
7	Quality Assurance Testing	0.300060	114.493	70.7535

Figure 6: Project Network Diagram of the Minimum Cost Solution for the Software Example with $T = 19$ and $Q_{LB} = 69$

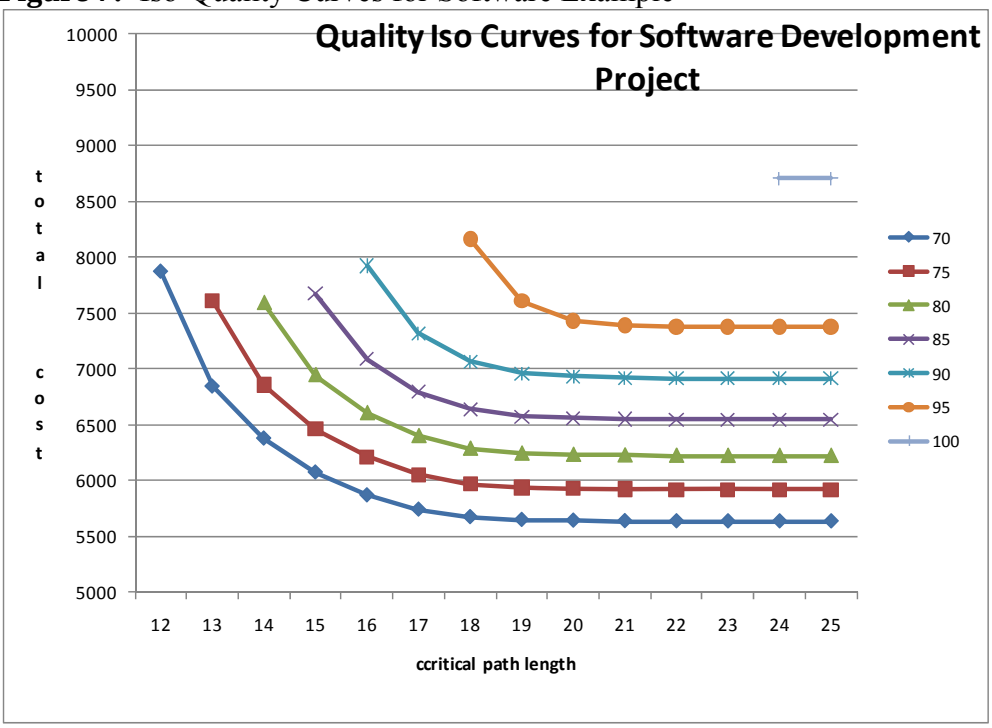


Iso-Curves

Figure 7 shows the iso-quality curves for the software development example. Assuming that the company wants at least a quality level of 90, they can inspect the curves for 90, 95, and 100, determine what seems optimal for each, and then apply judgment to decide whether or not the additional quality is worth the extra time and cost for the higher levels. In this case, at a quality level of 90, it seems the best time would be 19 weeks (as you reduce the time further, the cost increases much more significantly from that point), at a cost of a little under $7000. Khang and Myint [7] make a similar point in their analysis that for some quality levels, there seem to be budget threshold values of project times that are probably not worth the crashing effort. At 95, the best time seems to be 20 weeks, at a cost of a little under $7500. At 100, time would be 24 weeks, at a cost of about $8700. Depending on the circumstances, it does not seem that the added time and cost would be worth it to achieve the quality level of 100. On the other hand, it might well make sense to increase the quality from 90 to 95, if it only means taking 1 week

longer and costing $500 more. If desired, curves for 91 through 97 could also be generated for a finer analysis.

Figure 7: Iso-Quality Curves for Software Example



CONCLUSIONS AND FUTURE RESEARCH

The value of the project to the client can be measured in part by the level of quality associated with the completed project. Quality is acknowledged to be an important component of project management, but previously has received limited consideration in planning and scheduling. The standard implicit assumption is that time/cost trade-offs are evaluated given that some unspecified and unknown level of quality exists for each task. However, in many situations project managers must evaluate alternative options for accomplishing project activities, and these involve differing levels of time, cost, and quality. In such situations it makes sense to analyze the relationship between cost, time, and quality, and decide on their levels for each project task that best achieves the project's objectives. We have introduced the concept of a quality function that represents the relationships between quality, time, and cost for each task. Using two case studies, a translation agency and a software development company, the quality function is specified for each task and incorporated into a nonlinear programming model that allows quality to be explicitly considered in project planning and scheduling. An alternative formulation minimizes cost with bounds on project quality and completion time and leads to the creation of quality level curves. Both formulations can be very useful tools in making final project planning and scheduling decisions that explicitly model and incorporate quality.

Without explicitly considering quality, inferior decisions may be made concerning the planned levels of efforts for the various tasks. A key notion is that the quality of each project task is important and cannot be compromised without compromising the quality of the project itself. A

conscious part of the project planning process should include identification of alternative work plans that consider the tradeoffs between quality, time, and cost. If time and cost considerations are inflexible, the quality maximization model presented here can help project managers to select work plans for individual tasks that will increase overall project quality. If there is some flexibility in time and cost considerations, the iso-curve analysis can lead to important realizations such as that a significant increase in quality can occur with minimal impact on time and cost, or that time and/or cost could be appreciably improved with minimal effect on quality. The modeling framework proposed can provide insight on these issues and help to improve decision making in the project management planning and scheduling process.

An interesting insight is that by examining alternative optimal solutions additional quality for individual tasks can be obtained without additional cost. Once an optimal solution is obtained for either of our models, the solution can be checked to see if any tasks have slack and are not at their upper bound for time, since their times might be able to be increased (as long as the total project time never exceeds T_{UB}) without affecting the cost, and the quality of that task could be increased (as observed above in the translator $Q_{\min}$ example). If choices needed to be made between tasks whose time could be increased, those activities with the highest quality/time slope (slope of the curve sliced from quality surface by the vertical plane corresponding to the current cost for that task) would be increased first, analogous to what is done with cost in the traditional crashing problem. One way to achieve such a super-optimized solution mathematically would be to extend our problem formulation by adding a second-level goal (where maximizing $Q_{\min}$ or minimizing the cost was the top-level preemptive priority) of maximizing the *sum* of the qualities of all of the individual activities. We leave these extensions for future research.

Objective functions other than maximizing the minimum quality over all tasks are possible, such as maximizing average project quality or a weighted average of minimum and average project quality, or goal programming models using the three objectives. A consideration of these varying model objectives and their impact on optimal solutions can form the basis of future research.

Using the planned level of quality established for each task, future research can also investigate methods that can monitor and control quality during the course of the project, just as we now manage time and cost. Such methods will provide better balance in project management, emphasizing the nature and quality of the work completed, not just its schedule and budget.

REFERENCES

[1] H. Kerzner, *Project Management: A Systems Approach to Planning, Scheduling, and Controlling,* 8th ed. New York: Wiley, 2003.

[2] Project Management Institute, *A Guide to the Project Management Body of Knowledge*, 3rd ed., Newtown Square, PA, 2004.

[3] International Organization for Standards, ISO 9000:2000, 2000.

[4] B. Pollack-Johnson, and M. Liberatore, "Incorporating quality considerations into project time/cost trade-off analysis and decision making," *IEEE Transactions on Engineering Management*, vol. 53, no. 4, pp. 534 – 542, 2006.

[5] T. L. Saaty, *The Analytic Hierarchy Process*, Pittsburgh: RWS Publications, 1996.

[6] A. J. G. Babu and N. Suresh, N., "Project management with time, cost, and quality considerations," *European Journal of Operational Research*, vol. 88, no. 2, pp. 320-327, 1996.

[7] D. B. Khang and Y. M. Myint, "Time, cost and quality trade-off in project management: a case study," *International Journal of Project Management*, vol. 17, no. 4, pp. 249-256, 1999.

[8] L. O. Morgan, R. M. Morgan, and W. Moore, "Quality and time-to-market tradeoffs when there are multiple product generations," *Manufacturing & Services Operations Management*, vol. 3, no. 2, pp. 89-104, 2001.

[9] P. Brucker, A. Drexl, R. Mohring, K. Neumann, and E. Pesch, "Resource-constrained project scheduling: Notation, classification, models, and methods," *European Journal of Operational Research*, vol.112, no.1, pp. 3-41, 1999.

[10] Interagency Language Roundtable, "ILR skill level descriptions for translation performances," http://www.govtilr.org/Skills/AdoptedILRTranslationGuidelines.htm, Retrieved September, 16, 2011.

[11] American Translators Association, ATA Certification Program Rubric for Grading, Version 2009. Available from the American Translators Association, 225 Reinekers Lane, Suite 590, Alexandria, VA 22314.

[12] Lindo Systems Inc., *Extended Lingo/Win 32, Version 13.0*, Chicago, 2011.

PERSISTENCE IN COLLEGE MAJOR: A CLUSTER ANALYSIS APPROACH

John T. Quinn, Bryant University, (401) 232-6392, jquinn@bryant.edu
Phyllis A. Schumacher, Bryant University, (401) 232-6328, pschumac@bryant.edu
Alan D. Olinsky, Bryant University, (401) 232-6266, aolinsky@bryant.edu
Richard M. Smith, Bryant University, (401) 232-6064, rsmith@bryant.edu

ABSTRACT

Retention of students in their majors is an important concern for colleges and universities as it relates to planning issues, such as scheduling classes, particularly for upper classmen. This could also affect departmental requirements, like the size of the faculty. The math department has been collecting math placement scores and admissions data for all incoming freshmen at the university for many years. In the past, we have used this data mainly for placement in first-year classes and more recently to invite the most mathematically talented students to become mathematics majors. In this paper, we use the same data source to predict persistence in declared majors for all incoming students. In order to categorize the students, we use cluster analysis to see if students in particular majors share similar strengths based on the available data. Results are given in this paper.
Key Words: Cluster Analysis, Data Mining, Persistence in College Majors

INTRODUCTION

University academic departments are often concerned about the retention of students in their major. The necessity of scheduling classes, particularly for upper classmen, requires long-range planning. The math department has been collecting math placement scores and admissions data for all incoming freshmen at the university for many years. In the past, we have used this data mainly for placement in first-year classes and more recently to invite the most mathematically talented students to become mathematics majors. We have used regression, logistic regression, decision trees, and other data mining tools to predict which students would be successful mathematics majors (Smith and Schumacher (2005, 2006), Schumacher et al. (2010), Olinsky et al. (2011)). There have been numerous studies investigating the retention of students in majors, a few of which are listed here. For instance, Allen and Robbins (2008) used the generalized linear mixed model (GLMM) to analyze college major persistence considering interest-major fit and first-year academic performance work. Min *et al*. (2011) applied nonparametric survival analysis to look at the loss rate of undergraduate engineering students. Some of the factors involved in the study were SAT Math and Verbal scores, gender and ethnicity. Another paper that also researched engineering retention (first-year) is by Moses *et al*. (2011) which used binary logistic regression analysis, considering such variables as SAT

Math and Verbal scores and high school GPA. This study also looked at personality, based on scores from a five factor inventory.

We would now like to use the same data source to predict persistence in declared majors for all incoming students. In order to categorize the students, we have decided to use cluster analysis to see if students in particular majors share similar strengths based on the available data.

DATA CONSIDERATIONS

All 868 students who had made deposits for entrance in Fall 2007 were included in the original clustering in order to determine different profiles of incoming students. These students were chosen because we could then determine if they were persistent in their major through graduation. The variables used in the cluster analysis included verbal SAT (SAT_V), math SAT (SAT_M), High School Rank (Percentile), Score on Math Placement Exam (Score), and High School GPA (GPA). The Math Placement Exam score was based on the number of correct problems out of a possible 35. The High School Rank was recorded as a percentile in order to control for class size. Although the GPA scores did have values greater than 4, one of the GPA scores, which was erroneously recorded as a 400, was replaced by a missing value and thus a new variable which appears on SAS output as Replacement GPA is the GPA variable thus revised and now has a maximum value of 4.33. Finally, after these adjustments, all variables were then standardized in order for the units to become commensurate. This would avoid undue influence by the larger variables such as SAT scores which vary from 200 to 800.

CLUSTER ANALYSIS AND RESULTS

The cluster analysis was performed using SAS® Enterprise Miner. We utilized a k-means clustering algorithm in order to cluster the data into groups of students with similar academic characteristics. The algorithm determines the number of clusters by a cubic clustering criterion which minimizes the within cluster sum of squares. The procedure yielded six clusters, but since two of the clusters were very small, we opted to rerun the clustering algorithm and requested five clusters. This resulted in five distinct groups of students. The five clusters, which are shown in Figure 1, will be described below by the clustering variables in order of their inclusion in the cluster. The influence of each variable on its cluster is measured by the logworth of the factor. A picture of the SAS output illustrating the worth of the variables in each cluster is presented in Figure 2.

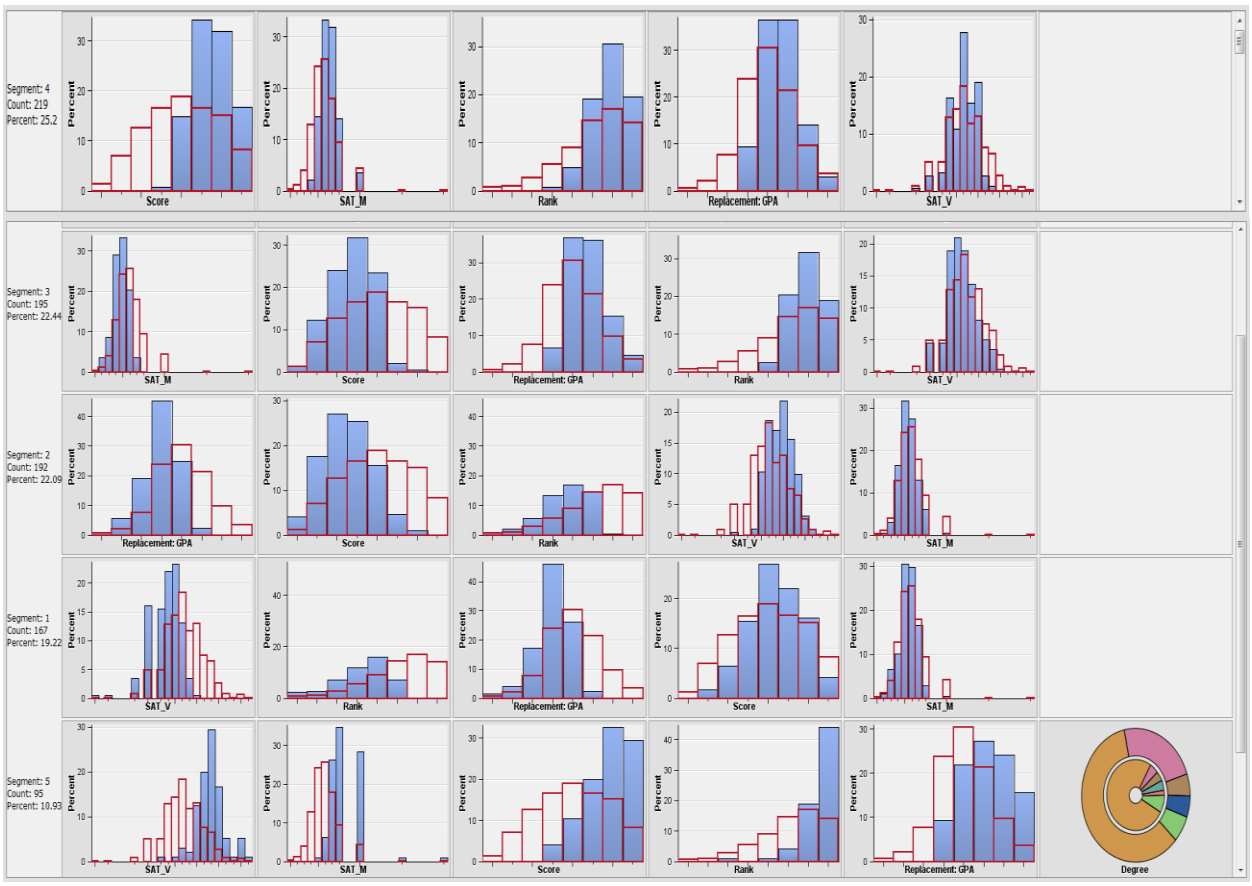


Figure 1: Clusters with Percentage of Students and Variables in Order of Discrimination

The largest cluster included 219 (25.2%) students and was driven by a high math placement score, which determined the cluster. The second determinant for this cluster was a high math SAT score. The third characteristic was a high rank in class, followed by a high GPA with the last variable being an average verbal SAT. The average values of the clustering variables for each cluster are given in Table 1. The second largest cluster included 195 (22.44%) students and was determined by low SAT_M scores, lower than average math placement scores, good GPA good rank in class, and average SAT_V. The third cluster included 192 (22.09%) students and was determined by average GPA, lower than average placement score, lower than average rank, high SAT_V, and average SAT_M. The fourth cluster included 167 (19.22%) students and was determined by lower than average SAT_V, lower than average rank, average GPA, average placement score and average SAT_M. The smallest cluster included 95 (10.93%) students and was determined by high SAT_V, high SAT_M, high placement score, high rank in class, and high GPA.

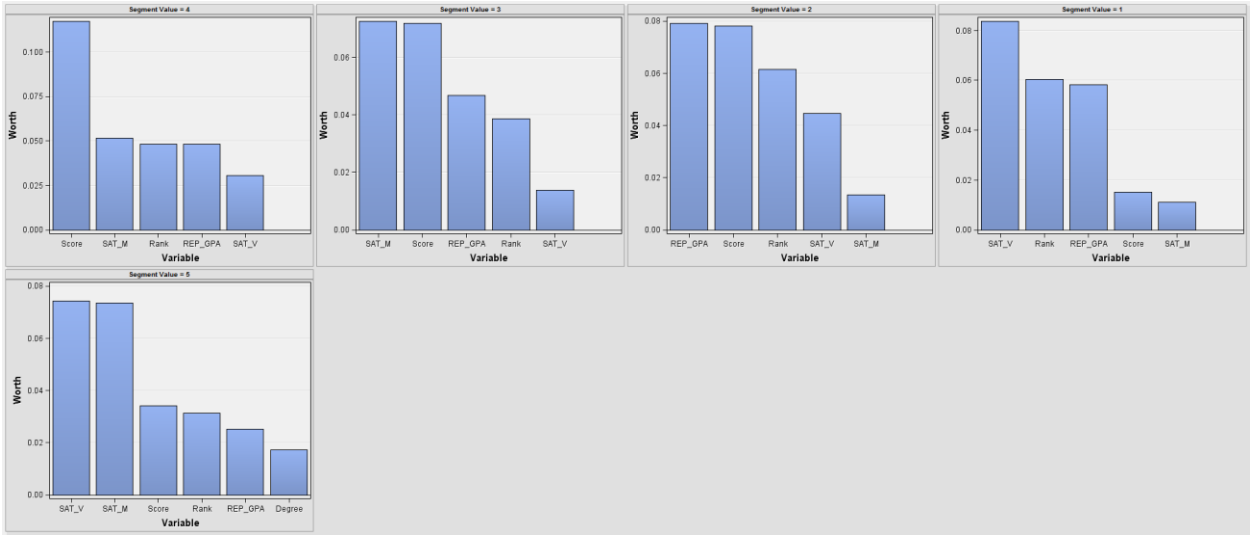


Figure 2: Clusters with Variables in Order of Discrimination and Measures of Worth

In reviewing these clusters, referring to lower than average as Low and higher than average as High, and combining math placement score with SAT_M and GPA with Rank, we have decided to classify the groups in order as: High Math/High Achieving/Average Verbal (HM/HA/AV), Low Math/High Achieving/Average Verbal (LM/HA/AV), Low Achieving/Low Math/High Verbal (LA/LM/HV), Low Verbal/Low Achieving/Average Math (LV/LA/AM), and High Verbal/High Math/High Achieving (HV/HM/HA). These groupings divide the students into distinct clusters which are recognizable. The first cluster of students, determined first by their high math scores and also being high achievers with average verbal scores, will most likely do very well in math and quantitative business areas. Students in the next two groupings, who have low math scores but either high achievement or high verbal skills, may excel in non-quantitative majors but probably not in the majors requiring good math skills. The fourth group, which was determined principally by their low verbal scores but also by low achievement and average math scores, will most likely need academic support to be successful in any major. The final group is determined primarily by their high verbal scores. They are also high achieving and have high scores in math and they should obviously excel in any major.

Cluster	Cluster Size	GPA	Rank	SAT_M	SAT_V	Score
(HM/HA/AV)	219	3.6	84.47	617.67	548.81	27.43
(LM/HA/AV)	195	3.63	85.13	542.17	542.63	16.55
(LA/LM/HV)	192	3.08	63.74	576.96	589.05	15.48
(LV/LA/AM)	167	3.1	59.15	574.06	496.35	22.81
(HV/HM/HA)	95	3.76	91.18	690.43	644.36	28.7

Table 1: Average Values of Clustering Variables by Cluster

Our intent in this paper is to determine the breakdown of clusters by majors for entering students, which clusters have the most students who persist in their chosen major and which majors have the most persistent students. We have recorded the majors the students selected as incoming

students in September of 2007. We also recorded the majors at graduation for those who graduated in May of 2011. Since our data was collected at Freshman Orientation in June and included all students who had submitted deposits, the students who did not graduate include those who never entered, those who transferred or dropped out, and those who took more than four years to graduate. Tables 2 and 3 indicate which majors were included in the different clusters both as incoming students and at graduation. An interesting observation can be made concerning the majors selected by the incoming students. As can be seen in Table 2, those who selected the math majors (AM-AMS) were mainly in the first and last clusters as is expected since they have high math scores. This is particularly of interest to us since we are in the Math Department. Another point is that, those who entered as undecided in Business (BUAD), were spread fairly uniformly throughout the five clusters (as a percentage of each clusters total), though more than half of the entering students in the LV/LA/AM cluster were undecided in business.

MAJ\CLUSTER	HM/HA/AV	LM/HA/AV	LA/LM/HV	LV/LA/AM	HV/HM/HA
ACG	31(14.16%)	18(9.23%)	14(7.29%)	9(5.39%)	7(7.37%)
AM-AMS	16(7.31%)	1(0.51%)	2(1.04%)	2(1.20%)	22(23.16%)
A&S OTHER	3(1.37%)	14(7.18%)	16(8.33%)	5(2.99%)	3(3.16%)
BUAD	91(41.55%)	76(38.97%)	75(39.06%)	87(52.10%)	34(35.79%)
COM	6(2.74%)	14(7.18%)	11(5.73%)	5(2.99%)	2(2.11%)
FIN	13(5.94%)	10(5.13%)	19(9.90%)	14(8.38%)	10(10.53%)
IB	20(9.13%)	16(8.21%)	12(6.25%)	14(8.38%)	6(6.32%)
LIBA	5(2.28%)	4(2.05%)	6(3.13%)	4(2.40%)	5(5.26%)
MGT-CIS-IT	20(9.13%)	22(11.28%)	16(8.33%)	13(7.8%)	3(3.16%)
MKT	14(6.93%)	20(10.26%)	21(10.94%)	14(8.38%)	3(3.16%)
TOTALS	**219(100%)**	**195(100%)**	**192(100%)**	**167(100%)**	**95(100%)**

Table 2: Majors of Entering Students by Clusters

In the same vein, with reference to table 3, it is interesting that of the three groups with low or average math scores, although some started in the math majors, there were no graduates in AM-AMS. Also of note from Table 3, the third and fourth groups, which are both low achievers, have the highest percentage of entering students who did not graduate in four years.

MAJ\CLUSTER	HM/HA/AV	LM/HA/AV	LA/LM/HV	LV/LA/AM	HV/HM/HA
ACG	47(21.46%)	31(15.90%)	23(11.98%)	18(10.78%)	13(13.68%)
AM-AMS	16(7.31%)	0(0.00%)	0(0.00%)	0(0.00%)	25(26.32%)
A & S OTHER	6(2.74%)	15(7.69%)	11(5.73%)	3(1.80%)	3(3.16%)
COM	6(2.74%)	11(5.64%)	11(5.73%)	6(3.59%)	3(3.16%)
FIN	37(16.89%)	22(11.28%)	29(15.10%)	23(13.77%)	13(13.68%)
IB	14(6.39%)	12(6.15%)	9(4.69%)	8(4.79%)	5(5.26%)
MGT-CIS-IT	21(9.59%)	30(15.38%)	22(11.46%)	24(14.37%)	9(9.47%)
MKT	28(12.79%)	39(20.00%)	25(32.02%)	25(14.97%)	6(6.32%)
Did Not Graduate	44(20.09%)	35(17.59%)	62(32.29%)	60(35.93%)	18(18.95%)
TOTALS	**219(100%)**	**195(100%)**	**192(100%)**	**167(100%)**	**95(100%)**

Table 3: Majors of Graduating Students by Cluster

Although students do not need to declare a major when they deposit, they must choose between the College of Business and the College of Arts and Sciences. Since many of the students were either undecided in the College of Business (BUAD) or undecided in the College of Arts and Science (LIBA), we then determined the breakdown of the clusters as to percentages of declared majors changing or not changing majors and for the undecided, who all choose a major by the beginning of junior year, whether or not they stayed in the same college. This information is presented in Table 4. It is of interest that the lowest achievers (LV/LA/AM) had the smallest percentage, 19.16%, who were not persistent in their chosen major. Also, the highest achieving group (HV/HM/AM) had the lowest percentage of (8.24%) changing majors.

Status\CLUSTER	HM/HA/AV	LM/HA/AV	LA/LM/HV	LV/LA/AM	HV/HM/HA
NO CHANGE	75(34.25%)	72(36.92%)	55(28.65%)	32(19.16%)	36(37.89%)
UND-SAME COLLEGE	59(26.94%)	60(30.77%)	45(23.44%)	50(29.94%)	25(26.32%)
CHANGE	29(13.24%)	23(11.79%)	22(11.46%)	20(11.98%)	8(8.24%)
UND-DIFF COLLEGE	12(5.48%)	5(2.56%)	8(4.17%)	5(2.99%)	8(8.42%)
DID NOT GRAD	44(20.09%)	35(17.95%)	62(32.29%)	60(35.93%)	18(18.95%)
TOTALS	**219(100%)**	**195(100%)**	**192(100%)**	**167(100%)**	**95(100%)**

Table 4: Students Changing/Not Changing Declared Majors and Undeclared Students Changing/ Not Changing Colleges Organized by Clusters

Finally in order to determine persistence in given majors, regardless of cluster, we include a table of those who started in a particular major and graduated in the same major, changed to a different major, or did not graduate, regardless of college. These results appear in Table 5. With reference to the data in table 5, it is notable that the Accounting major had the greatest percentage of persistence in major (69.62%) and this closely followed by AM-AMS (62.8%) and

Communication (60.5%). Generally, it is true that for those entering with a declared major, for the most part close to 50% persisted in their majors. The only majors significantly below 50% (34.2%) were those grouped together as Other A&S. These students were in very small majors without many fellow students in the same disciplines, which could have affected their decisions to switch. This deserves further study.

Major/Status	No Change	Change	Did Not Graduate
AM-AMS(43)	62.8%	18.6%	18.6%
COM(38)	60.5%	21%	18.5%
OTHER A&S(41)	34.2%	34.1%	31.7%
ACG(79)	69.62%	13.92%	16.46%
FIN(66)	57.6%	19.7%	22.7%
IB(68)	47%	20.6%	32.4%
MKT(72)	55.6%	20.8%	23.6%
MGT-CIS-IT(74)	52.7%	25.7%	21.6%

Table 5: Students Changing/Not Changing Declared Majors Organized by Major

DISCUSSION

It is our intention to use the results of this study to assist us in initial counseling of our freshmen students. For example, the fact that the students who can be identified as low achievers, are indeed less likely to graduate and thus should be counseled to take advantage of the many student support services available to them. We in the math department can also counsel students who are unlikely to succeed by noting that although a few students started in one of the three clusters without high math scores, there were no students from these clusters who graduated as a math major. This information is especially useful as many of our students enter with undeclared majors. Hopefully, our clustering results, and examination of persistence of majors in each cluster through graduation, will allow us to improve such counseling of incoming freshmen and also assist us in the proper placement of students into appropriate freshman courses. In addition, it will help us to develop a long-range schedule of upper-level courses in the department. Finally, we feel that this is an interesting, and somewhat unique, application of cluster analysis in higher education.

REFERENCES

Allen, J. & Robbins, S. B., "Prediction of College Major Persistence Based on Vocational Interests, Academic Preparation, and First-Year Academic Performance," *Research in Higher Education*, Vol. 49 (2008), pp. 62-79.

Min, Y., Zhang, G., Long, R. A., Anderson, T. J., & Ohland, M. W., "Nonparametric Survival Analysis of the Loss Rate of Undergraduate Engineering Students." *Journal of Engineering Education*, Vol. 100 (2011), No. 2, pp. 349-373.

Moses, L., Hall, C., Wuensch, K., De Urquidi, K., Kauffmann, P., Swart, W., Duncan, S., & Dixon, G., "Are Math Readiness and Personality Predictive of First-Year Retention in Engineering?" *The Journal of Psychology*, Vol. 145 (2011), Iss. 3, pp. 229-245.

Olinsky, A., Schumacher, P., & Quinn, J.T., "Assessing Data Mining Approaches for Analyzing Actuarial Student Success Rate." In Zhang, Q., Segall, R., & Cao, M (Eds.) in *Visual Analytics and Interactive Technologies: Data, Text and Web Mining Applications*, Hershey, Pennsylvania: IGI Global Publishing. (2011)

SAS Enterprise Miner (version 6.1), SAS Corporation, Cary, NC (2011).

Schumacher, P., Olinsky, A., Quinn, J.T., & Smith, R.M.. "A Comparison of Logistic Regression, Neural Networks, and Classification Trees Predicting Success of Actuarial Students," *Journal of Education for Business* , Vol. 85 (2010), Number 5, pp. 258-263.

Smith R. M., Schumacher P. "Academic Attributes of College Freshmen which Lead to Success in Actuarial Studies in a Business College." *Journal of Education for Business*, Vol. 81(2006), Number 5, pp. 256-260.

Smith R. M., Schumacher P. "Predicting Success for Actuarial Students in Undergraduate Mathematics Courses." *College Student Journal*, Vol. 39 (2005), Number 1, pp.165-178.

WHAT DETERMINES BOX OFFICE SUCCESS OF A MOVIE IN THE UNITED STATES?

Borga Deniz
Christopher Newport University, Luter School of Business, 1 University Place, Newport News, VA 23606
Phone: (757) 594 8915, e-mail: borga.deniz@cnu.edu

Robert B. Hasbrouck
Christopher Newport University, Luter School of Business, 1 University Place, Newport News, VA 23606
Phone: (757) 594 7174, e-mail: rhasbro@cnu.edu

ABSTRACT

Motion picture business in the United States is a multi-billion dollar industry which is an important part of the country's economy. There were over 120,000 movies that were shown in movie theaters across the U.S. in 2010. Some are successful and many are not. What determines box office success of a movie in the United States? In this paper we investigate the impact of genre, MPAA rating, budget, star power, adaptation from another medium , sequels and remakes on total U.S. box-office revenue. We use data from one hundred and fifty top grossing movies of 2010, and we present our multiple linear regression analysis results.

INTRODUCTION AND LITERATURE REVIEW

The movie industry in the United States has a tremendous importance for the United States economy. In the U.S., movies made more than 10 billion dollars in 2010 and the U.S. movie industry employs more than half a million people. Movie industry aims to entertain millions of viewers and it depends on the preferences of the movie-goers. If viewers do not choose to see a movie millions of dollars can be lost, companies and producers can go bankrupt; therefore we can say that entertainment business is a serious business. For production companies it is very important to predict what movie is going to be a success or a flop. This paper investigates the determinant factors on a movie's success.

In the literature there are many research papers that explore the determinants of motion picture box office revenue. Litman (1983) was the first to develop a multiple regression model to predict the commercial success of movies [7]. The independent variables in this work were movie genre (science fiction, drama, action-adventure, comedy, and musical), Motion Picture Association of America (MPAA) rating (G, PG, R and X), superstar in the cast, production costs, release company (major or independent), Academy Awards (nominations and winning in a major category), and release date (Christmas, Memorial Day, Summer). Litman's model provides evidence that the independent variables of production costs, critics' ratings, science fiction genre, major distributor, Christmas release, Academy Award nomination, and winning an Academy Award are all significant in the success of a film. Litman and Kohl (1989) [9], Litman and Ahn (1998) [8], Terry et al. (2004) [18] and Brewer et al. (2009) [3] have replicated and expanded the initial work of Litman. Terry et al. (2009) did a similar analysis on English language movies in terms of their foreign box office revenue [20]. Terry and De'Armond (2008) analyzed determinants of movie video rental revenue [19]. Smith and Smith (1986) is another early study to examine the performance of movies [16]. They analyzed the determinants of successful films, as defined by all-time box-office revenues. They observed that movie have become increasingly more specialized as a result of television. Prag and Casavant (1994) found a positive impact of star power, praise by critics, sequels and Academy Awards on revenues when there is no marketing expenditures [11]. Star power, Academy Awards, and production costs are found to be positive determinants of advertising spending. One area of interest in the literature has been the role of critics (Weiman, 1991) [21]. The majority of the literature finds that critics play a significant role on the success of a movie. According to Eliashberg and Shugan (1997) [5] there are two types of critics: the influencer and the predictor. The influencer is a critic that will influence the box office results of a film based on his or her review. Eliashberg and Shugan's results suggest that critics do have the ability to manipulate box office revenues based on their review. The predictor predicts the success of a movie but the review will not necessarily have an impact on how well the movie performs commercially. Eliashberg and Shugan find that the predictor role of a critic is statistically less important than the influencer role. King (2007) [6] also explores the power of critics on revenue of movies. He concludes that there is no correlation between critical ratings for movies and their commercial success when all releases are considered because of the affinity most critics have for foreign movies and documentaries compared to the average movie-goer, and if one considers only the films released to a wide audience (more than 1,000 theaters) then critical ratings have a significant positive impact on revenue. Reinstein and Snyder (2000) investigate impact of the critics Siskel and Ebert's reviews on commercial success [14]. They conclude that positive reviews have a large impact on box office success. Reinstein and Snyder also report that entire critic population' influence on box office is not necessarily significant but only a few critics' reviews can influence a movie's revenue significantly.

Research has also shown a there is significant relation between season of film's releases and its revenue. Litman (1983) indicated that the most important time for a film release is the Christmas season [7]. However Sochay (1994) reported that the summer is the best season to release a movie [17]. Sochay, referencing Litman (1983), explains the conflict in these two results is based on competition. Sochay mentions that the successful season can shift from the summer to Christmas from year to year based on film distributors' effort to avoid strong competition.

Motion Picture Association of America (MPAA) ratings may also influence box office revenue of a movie. Movie production companies usually try hard to get a better rating for their movies. To that end they frequently reshoot or reedit scenes numerous times in order to get their preferred ratings which are usually the coveted PG or PG-13 ratings. These two ratings are the best ratings for producers as, practically, they will not keep anyone from seeing the movie. Anast (1967) was the first to study the relationship between film genre and movie attendance [1]. His results showed that films with violence and eroticism had a positive correlation while action-adventure genre had negative correlation with movie attendance. Litman (1983) concluded that that film ratings do not have significant effect on a film's box office success unless the movie's genre is science fiction [7]. Austin (1984) also looked at film ratings to see if there is a correlation between ratings and movie attendance but could not find a significant relation [2]. However Ravid (1999) showed that G and PG rated films have a positive impact in the box office [12]. Furthermore Terry et al. (2004) found that the negative effect of the R rating on box office revenue is in the amount of $10 million on average [18].

In movie industry awards are very important as they are highly publicized in the media. Commercial effect of an award is first investigated by Litman (1983) [7]. He found that an Academy Award nomination in the categories of best actor, best actress, and best picture is worth $7.34 million, while a win in one of these major categories can translate into over $16 million at the box office. Nelson et al.(2001) [10] estimated that an Academy Award nomination in a major category is worth $4.8 million and a victory brings in $12 million on average. They indicate that in the movie industry it is a common practice to delay film releases toward the end of the year as it improves the chances of receiving nominations and monetary rewards. Dodds and Holbrook (1988) look at the impact of an Academy Award after the nominations have been announced and after the award ceremony [4]. The authors find that a nomination for best actor is worth about $6.5 million, best actress is worth $7 million and best picture is worth $7.9 million. After the award ceremony the best actor award is worth $8.3 million, best picture is worth $27 million, and best actress award is not statistically significant. Simonoff and Sparrow (2000) find that for a movie opening on less than ten screens, an Academy Award nomination will increase the movies expected gross close to 250% more than it would have grossed if it had not received the nomination. For movies opening on more than ten screens, an Academy Award nomination will increase the movies gross by nearly 30%. [15]

Ravid (1999) examined film revenue and return-on- investment (ROI) as functions of production cost and star actors, among other variables [12]. Regressions show that large production costs significantly increase film revenue, but do not increase the ROI. The quantity of critic reviews, which is representative of exposure and availability to the public, is positively significant. Sequels are found to perform significantly better than non-sequels. While univariate tests suggest that movie stars increase revenue, regressions find star presence to be insignificant. This supports the 'rent-capture' hypothesis that stars earn salaries equivalent to their market value and do not impact the profitability of films. Ravid (2004) emphasized the risk associated with making movies by estimating a ROI model [13]. The study focuses on the strategies utilized by studio executives when choosing the films to be released.

Unlike many other studies, the effect of violence in R-rated films is examined. The results illustrate that high-violence films are expected to be financially 'safer' to produce. On average, the ROI of violent films falls within the middle of the sample's distribution.

DATA, MODEL and ANALYSIS

In movie theaters across the United States 123,340 films were shown in 2010 and these movies approximately grossed $10.5 billion in the box office. Our data set in this study is composed of 150 top grossing movies released in the year 2010. The movies in the sample include a range of movies from *Toy Story 3* (the top grossing movie of 2010) with $415 million to *Buried* (the 150th movie in our sample) with 1.03 million dollars US box office revenue. These top 150 movies grossed about 9.8 billion dollars which makes up around 94% of the total domestic box office revenue for movies of 2010.

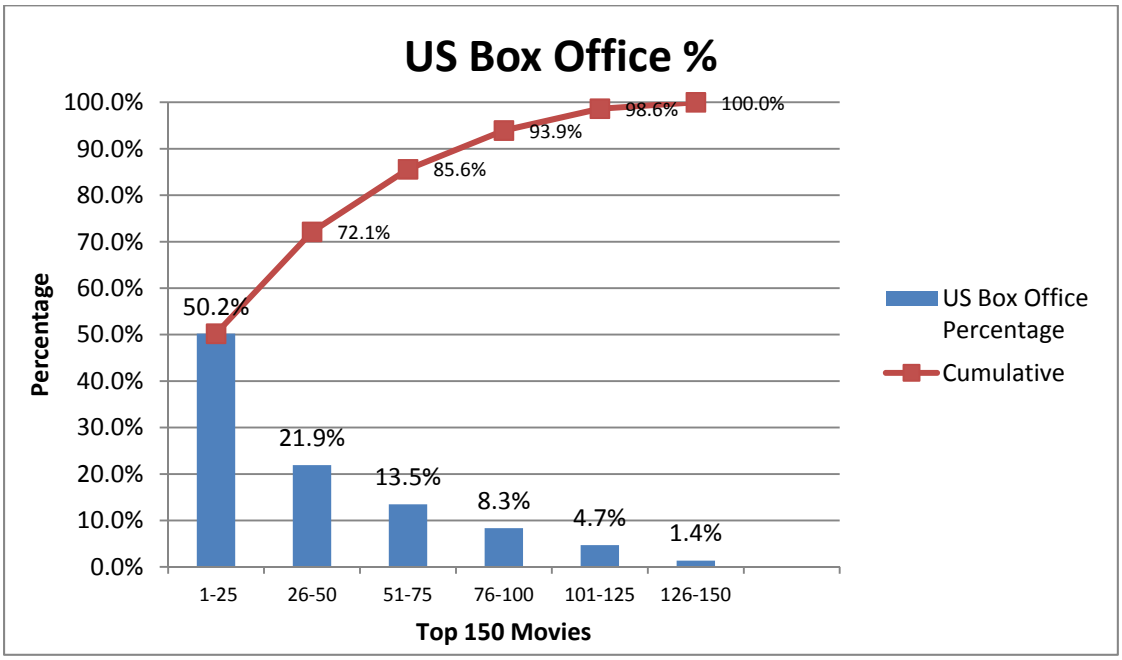


Figure-1: Approximately 50% of the box revenue comes from the top 25 movies.

Average revenue for a top-150 movie, in million dollars, is 65.4 with a standard deviation of 74. A remake makes 65.7 million on average. If a movie is a sequel then it averages 135 million; however a non-sequel averages 54.7 million. Movies with at least one *star* in them average 87.3 million; a movie without a star makes 45.2 million on average. (As a general rule we define a movie star as follows: A movie star is an actor/actress who has been paid at least $10M for a movie role in his/her career.) The Motion Picture Association of America (MPAA) ratings are G (general), PG (parental guidance suggested/some material might not be suitable for children), PG-13 (parents strongly cautioned/some material may be inappropriate for children under the age of 13), R (Restricted / Children Under 17 Require Accompanying Parent or Adult Guardian.), and NC-17 (No One 17 and Under Admitted). A movie with an R rating averages 39.2 million, a PG-13 movie makes 71.8 million, and movie with a PG or G rating earns 105 million on average. This largely explains why studios try hard to get a lower rating for their movie from MPAA.

Genre of movie can also be an important factor on its box office success. Examples of movie genre are: Comedy, Sci-Fi, Horror, Action, Romance, Drama, Adventure, Fantasy, Family, Crime, Thriller, Mystery, Musical, Crime. A movie can be in more than one single genre, like *romantic comedies* or *action comedies*. In our data genre of a movie is identified according the popular movie website, imdb.com. Based on our study we can say that most bankable movie genre is horror: On average a horror movie makes 7.27 times its production cost. For example the production cost of *Paranormal Activity 2* was only 2.8 million while it made 84.7 million dollars (30.8 times its cost) in the box office.

The following table shows the variables we use in our study:

VARIABLE	DEFINITION
USBoxOffice	US box office earnings of a movie in millions of dollars
Revenue/Budget	Ratio of USBoxOffice to Budget
Budget	Production cost of a movie
Sequel	categorical variable for movies that follow a previously released film
Star Power	categorical variable for films that have a *movie star* in a leading role
Remake	categorical variable for movies that are remake of another film
Adaptation	categorical variable for movies that are adapted from another medium
R	categorical variable for movies that are rated R
PG-13	categorical variable for movies that are rated PG-13
G	categorical variable for movies that are rated G
PG	categorical variable for movies that are rated PG
PG or G	categorical variable for movies that are rated PG or G
Comedy	categorical variable for movies that can be categorized as comedy
Sci-Fi	categorical variable for movies that are science fiction
Horror	categorical variable for horror movies
Action	categorical variable for action movies
Romance	categorical variable for romantic movies
Drama	categorical variable for dramas
Adventure	categorical variable for adventure genre
Fantasy	categorical variable for fantasy movies
Family	categorical variable for family movies
Crime	categorical variable for crime movies
Thriller	categorical variable for thrillers
Mystery	categorical variable for mysteries
Musical	categorical variable for musicals
Crime	categorical variable for crime movies
War	categorical variable for war movies
Biography	categorical variable for biographies
Western	categorical variable for westerns
Documentary	categorical variable for documentaries

Table-1: Definition of variables

In the following table we present correlation coefficient for all input variables with US Box Office and whether they are significant as a single variable. As it is seen below, Budget correlates with USBoxOffice the highest. R and Drama negatively correlate with the dependent variable. Apparently Horror, Sport, Romance, Musical, PG-13, Crime, War, Biography,

Comedy, Adaptation, Thriller, Mystery, Western, Documentary and Remake are not significant variables.

	Corr. Coeff. with US Box Office	p-value	Significant (α=.05)
Budget	0.706	0	YES
Adventure	0.528	0	YES
Animation	0.405	0	YES
Fantasy	0.387	0	YES
Sequel	0.37	0	YES
Family	0.34	0	YES
R	-0.294	0	YES
Star Power	0.285	0	YES
PG or G	0.274	0	YES
G	0.245	0.001	YES
Drama	-0.227	0.003	YES
PG	0.21	0.005	YES
Action	0.18	0.014	YES
Sci-Fi	0.154	0.03	YES
Horror	-0.111	0.089	NO
Sport	0.087	0.146	NO
Romance	-0.087	0.146	NO
Musical	0.086	0.147	NO
PG-13	0.069	0.202	NO
Crime	-0.065	0.213	NO
War	-0.048	0.279	NO
Biography	-0.039	0.319	NO
Comedy	0.025	0.382	NO
Adaptation	-0.025	0.382	NO
Thriller	-0.02	0.405	NO
Mystery	-0.013	0.439	NO
Western	0.011	0.445	NO
Documentary	-0.008	0.464	NO
Remake	0.001	0.493	NO

Table-2: Correlation of variables with US Box Office

The following three tables below depict the linear regression model that fits the data best when USBoxOffice is the dependant variable. Several input variables considered originally for the model are excluded because of problems such as statistical significance and multi-collinearity concerns.

	Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
	B	Std. Error	Beta	t	Sig.	Tolerance	VIF
(Constant)	3.884	6.141		0.632	0.528		
Budget	0.641	0.099	0.454	6.481	0	0.549	1.82
Sequel	46.845	11.792	0.216	3.973	0	0.914	1.094
Animation	52.205	17.83	0.168	2.928	0.004	0.82	1.22
Sport	70.194	27.532	0.133	2.55	0.012	0.989	1.011
Adventure	30.82	11.508	0.173	2.678	0.008	0.647	1.547
Star Power	21.968	8.221	0.149	2.672	0.008	0.871	1.148

Table-3: Model coefficients (dependent variable: USBoxOffice)

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	500905.487	6	83484.248	37.877	0
Residual	315180.802	143	2204.062		
Total	816086.289	149			

Table-4: ANOVA Table (dependent variable: USBoxOffice)

Model Summary				
R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
0.783	0.614	0.598	46.947	1.37

Predictors: (Constant), Adventure, Star Power, Sport, Sequel, Animation, Budget

Table-5: Model summary (dependent variable: USBoxOffice)

The model explains sixty percent of the variation in the domestic box office revenue. According to the model: Having Adventure as a genre adds $31 million to the box office earnings while Star Power adds $22 million. Having Sport as genre contributes $70 million, having animation contributes $52 million to the box office. Being a sequel improves box office success by $47 million.

Alternatively we also used the ratio of revenue-to-budget of a film as a dependent variable. In the following table we present correlation coefficients for all input variables with Revenue/Budget and whether they are significant as a single variable.

	Correlation w/ Revenue to Budget ratio	p-value	Significant (α=.05)
Horror	0.385	0.000	YES
Budget	-0.2	0.007	YES
Adaptation	-0.164	0.023	YES
Star Power	-0.163	0.023	YES
Mystery	0.128	0.059	NO
Action	-0.124	0.066	NO
Comedy	-0.114	0.082	NO
Sequel	0.102	0.107	NO
Fantasy	-0.095	0.124	NO
Romance	-0.09	0.138	NO
PG	-0.087	0.145	NO
PG or G	-0.087	0.146	NO
Adventure	-0.087	0.145	NO
PG-13	0.086	0.147	NO
Thriller	0.086	0.148	NO
Family	-0.056	0.248	NO
Sci-Fi	-0.056	0.247	NO
Drama	-0.055	0.25	NO
Crime	-0.054	0.258	NO
Remake	-0.051	0.269	NO
Documentary	0.051	0.268	NO
Sport	0.039	0.317	NO
Musical	-0.037	0.326	NO
Animation	-0.036	0.333	NO
Western	-0.023	0.388	NO
Biography	0.022	0.395	NO
War	-0.02	0.404	NO
R	-0.014	0.431	NO
G	-0.006	0.473	NO

Table-6: Correlation of variables with Revenue/Budget

As it can seen above Horror is the only significant variable that positively correlates with Revenue/Budget. It is consistent with the fact that horror movies usually do not have stars in them, which increases costs significantly. The model that best fits the data is presented as follows:

Coefficients (Dependent Variable: Revenue/Budget)								
		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
	(Constant)	1.801	0.35		5.153	0		
	Horror	6.517	1.159	0.441	5.62	0	0.907	1.102
	Remake	-2.448	1.038	-0.185	-2.359	0.02	0.907	1.102

Table-7: Model coefficients (dependent variable: Revenue/Budget)

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	496.143	2	248.071	16.025	0
Residual	2275.63	147	15.48		
Total	2771.772	149			

Table-8: ANOVA Table (dependent variable: Revenue/Budget)

Model Summary				
R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
0.423	0.179	0.168	3.934523	2.005

Predictors: (Constant), Horror, Remake

Table-9: Model summary (dependent variable: Revenue/Budget)

This model explains seventeen percent of the variation in Revenue/Budget variable. According to the model a Horror movie contributes to the Revenue/Budget ratio by 6.5 but a remake negatively impacts the ratio. Horror movies usually have low budgets but they generate strong revenue relative to their cost. Horror movie audience is usually young, story and special effects are more important than seeing a Hollywood star in a movie (e.g. *Paranormal Activity* and *Saw* series, *Insidious*, *The Last Exorcism*). Advances in digital film making made horror movies an easy entry genre as their production cost decreased due to technology.

CONCLUSION AND RESEARCH DIRECTIONS

In terms of domestic box office success Sequel, Animation, Adventure, Star Power and Budget have statistically significant and positive effect based on our study. For example Star Power adds about $22 million, Adventure adds $31 million to the revenue. In our study we also observed that, in terms of revenue-to-budget ratio the most successful genre is horror. In terms of future research one can study worldwide box office revenue of movies. Another research extension can be to expand the domestic (or worldwide) revenue including earnings from DVD/Blu-Ray rentals (and sales), pay-per-view and TV.

REFERENCES

[1] Anast, Philip (1967). Differential Movie Appeals as Correlates of Attendance. *Journalism Quarterly*, 44, 86-90.

[2] Austin, Bruce A. (1984). Portrait of an Art Film Audience. *Journal of Communication*, 34 (winter), 74-87.

[3] Brewer, S. M. & J. M. Kelley & J. J. Jozefowicz, 2009. "A blueprint for success in the US film industry," *Applied Economics*, Taylor and Francis Journals, vol. 41(5), pages 589-606.

[4] Dodds, John C. & Morris B. Holbrook (1988). What's an Oscar Worth? An Empirical Estimation of the Effect of Nominations and Awards on Movie Distribution and Revenues. *Current Research in Film: Audiences, Economics and the Law*, Vol. 4.

[5] Eliashberg, J. and Shugan, S. M. (1997) Film critics: influencers or predictors? *Journal of Marketing*, 61, 68–78.

[6] King, Timothy (2007). Does film criticism affect box office earnings? Evidence from movies released in the U.S. in 2003. *Journal of Cultural Economics*, 31, 171-186.

[7] Litman, Barry R. (1983). Predicting Success of Theatrical Movies: An Empirical Study. *Journal of Popular Culture*, 16 (spring), 159-175.

[8] Litman, Barry R. & H. Ahn (1998). Predicting Financial Success of Motion Pictures. In B.R. Litman, *The Motion Picture Mega-Industry*, Allyn & Bacon Publishing, Inc.: Boston, MA.

[9] Litman, Barry R. & A. Kohl (1989). Predicting Financial Success of Motion Pictures: The 80's Experience. *The Journal of Media Economics*, 2 (1), 35-50.

[10] Nelson, Randy A., Michael R. Donihue, Donald M. Waldman & Calbraith Wheaton (2001). What's an Oscar Worth? *Economic Inquiry*, 39 (1), 1-16.

[11] Prag, J. and Casavant, J. (1994) An empirical study of the determinants of revenues and marketing expenditures in the motion picture industry, *Journal of Cultural Economics*, 18, 217–35.

[12] Ravid, S. A. (1999) Information, blockbusters, and stars: a study of the film industry, *The Journal of Business*, 4, 463–92.

[13] Ravid, S. A. (2004) Are they all crazy or just risk averse? Some movie puzzles and possible solutions, in *Contributions to Economic Analysis*, Vol. 260 (Ed.) V. A. Ginsburgh, Elsevier, New York, NY, pp. 33–47.

[14] Reinstein, David A. & Christopher M. Snyder (2000). The Influence of Expert Reviews on Consumer Demand for Experience Goods: A Case Study of Movie Critics. Working Paper, University of California-Berkeley and George Washington University.

[15] Simonoff, Jeffrey S. & Ilana R. Sparrow (2000). Predicting Movie Grosses: Winners and Losers, Blockbusters and Sleepers. *Chance Magazine*, 13 (3), 15-24.

[16] Smith, Sharon P. & V. Kerry Smith (1986). Successful Movies: A Preliminary Empirical Analysis. *Applied Economics*, 18, 501-507.

[17] Sochay, Scott (1994). Predicting the Performance of Motion Pictures. *The Journal of Media Economics*, 7 (4), 1-20.

[18] Terry, Neil, Michael Butler & De'Arno De'Armond (2004). The Economic Impact of Movie Critics on Box Office Performance. *Academy of Marketing Studies Journal*, 8 (1), page 61-73.

[19] Terry, Neil & De'Arno De'Armond (2008). The Determinants of Movie Rental Revenue Earnings. *Academy of Marketing Studies Journal*, 12 (2), page 35-47.

[20] Terry, N., Cooley, J. W., and Zachary, M. A. (2009). "The determinants of foreign box office revenue for english language movies." *Journal of International Business and Cultural Studies*, 2(1).

[21] Weiman, Gabriel (1991). The Influentials: Back to the Concept of Opinion Leaders. *Public Opinion Quarterly*, 55, 267-279.

THE CHALLENGES OF REAL WORLD ANALYTICS: A CASE STUDY FOR THE CLASSROOM

Michael Salzillo, Bryant University, (401) 232-6897, msalzill@bryant.edu
Kristin Kennedy, Bryant University, (401) 232-6316, kkennedy@bryant.edu
Alan Olinsky, Bryant University, (401) 232-6266, aolinsky@bryant.edu

ABSTRACT

The transition from having students work with textbook data to actual raw data is monumental. The authors of this paper were given the opportunity to work with the Providence Water Authority (PWA) in Providence, Rhode Island. The PWA had a very large data set, and they wanted help with their statistical analysis. This paper will highlight issues we faced in cleaning and grappling with the data. In addition to consulting with the PWA, we now plan on using this dataset as a case study in an upcoming analytics course.

INTRODUCTION

In teaching statistics, especially in upper level university classes, such as junior/senior classes, and most importantly in graduate classes, it is crucially important to let the students work with real data and large data sets. The transition from allowing students to work with textbook data to actual raw data is monumental. However, professors know that introducing students to real data sets will often, if not always, result in difficulties with the data. This can be especially true, if the data comes from a company that would like some (or any) data analysis performed free of charge. Without careful review and preparation of the data submitted by a company for some review, it can take most of a semester to simply get the data into the correct format for analysis. However, this is also an important step that cannot be overlooked or glanced over. All students need to encounter the difficulties of "cleaning the data" especially in higher level statistics classes.

If there are no large data sets to analyze for a classroom situation, data can be downloaded from the internet. It can be more interesting for the students, though, if they work with a large data set from a company struggling with real issues. But professors beware! It is often true that the management of the company not only has the data in very poor format, but they also do not have an exact goal or question that they would like to see resolved. This is a twofold problem. Not only is cleaning the data imperative, but professors must begin to make some overall sense of the

data and formulate possible questions that could be answered using the data. First there must be a clear understanding of the data, before one can formulate questions to be answered using the data. This seems straightforward, but it is quite time-consuming.

Literature Review

As mentioned in an article by Tweedie and Taylor, the ability to understanding our data and the questions of the client are most important. Close examination can reveal far more than the client originally told us. [1] This was expecially true in working with this case.

As stated by Zahn and Isenberg, "…statistical consulting is a complex activity that requires statistical and nonstatistical skills. Together, these skills determine the ultimate success of the consultation." They further discuss a model that consists of 4 parts: "(a) identification of relevant aspects of the problem situation, (b) definition of the client's goals, (c) determination of the actions to be taken, (d) discussion of various aspects of the consulting relationship and who will do what when." They indicate that the consultant needs to have various nonstatistical skills in order to deal with these issues. [2] An article by Kirk discusses coping with clients having diverse statistical backgrounds and dispelling inappropriate expectations by clients. [3]

A comprehensive presentation on data cleaning is presented in a special edition of Data Engineering (December 2000). In one of the articles, Rahm and Do classify data quality problems that are addressed by data cleaning and provide an overview of the main solution approaches. [4]

The authors of this paper were given the opportunity to work with the Providence Water Authority (PWA) in Providence, Rhode Island. The PWA had a very large data set, and they wanted help in statistical analysis. This paper will highlight issues we faced in cleaning and grappling with the data. In the end, some of the questions that they were initially asking were simply unanswerable with the given data set. In fact, some of their initial questions were vague, and we needed to carefully identify questions that could be answered once the data was in good form. This example highlights that there is a continual conversation between analyst and client that takes time. The analyst must learn and understand first what the important variables are within a set and second what specific questions can actually be asked of the data. This is a process that the analyst goes through with little help from the client, and the client can be surprised at some of the difficulties encountered with the data. Throughout the process the authors used Enterprise Guide [5] for all analysis. We found this package very easy to use, and it handled the large data set quickly.

Background Information and Project Data

The PWA supervises the water flow to the public in the cities of Providence, North Providence, Cranston, Johnston, and Warwick, all in Rhode Island. The organization, which is governed by the Providence City Council, was instituted in 1773 and has gone through many changes as the State grew in population over many years. Today the main water supply is the Scituate Reservoir, and the treatment plant for this supply also resides in the town of Scituate. The plant has a maximum water capacity of 144 million gallons of water flow per day and still remains the largest treatment facility in New England. The plant supplies water to both approximately 72,300 retail service connections and to 8 wholesale communities through a system of 870 water mains [6].

The PWA has recently been approached by government and environmental agencies to improve their pricing structure and, if possible, show a higher profit. They had a large data set dating back to 2006, and management of the PWA initially hoped to identify customers who were trying to conserve water consumption, an admirable goal.

The PWA had collected data from 2006 – 2011, but it was only partially collected in 2006 and 2011. The set of data had 1.5 million observations. Each observation had 10 fields. For example, some of the fields were: a service location, billing address, cubic flow for a particular reading, date of the reading, a dwelling code, and some other less important fields. The main variable that we wanted to examine was the cubic flow, which could be translated into a consumption rate.

Readings are taken more or less every quarter. They hope to increase readings to every month in the future. A drive by electronic reading is taken on all meters at each service location, approximately every three months.

The analysis of consumption rates would certainly depend on the seasons. We assigned a seasonal label for each reading, so that we could compare each season for different years for any location. We presumed that for residential readings, more water is used during summer months as opposed to winter months. It would be expected that commercial usage would remain steady throughout the year and recreational usage, such as parks or beaches, would only show a significant flow during the summer months.

It took the better part of a year to clean the data and represent the data in a format that was understandable enough so that some data analysis could begin.

DATA CLEANSING

The first problem encountered was trying to clarify the objectives of the PWA. Generally, they hoped that we could identify residential homes or single businesses that were practicing good consumption. Secondly, we needed an understanding of the meaning of all the fields. For example, it was confusing having a service location and an account number for each observation. In many cases these two were the same, but that was not always true. We asked that the PWA give us any reports they might have, which would have been generated by them in the past for their own use. These reports did help us to clearly identify what the meaning of every field was. A customer could have one account number, with several service location numbers.

There were technical issues to clarify. Originally we were given 1.5 million records in two separate Excel spread sheets. Excel cannot append these two files into one. We then choose to use SAS Enterprise Guide for any analysis, and we were able to append the two files in SAS.

There were problems with partial data for certain years. The data did range in date from 2006 to 2011, but there were only partial readings for 2006 and 2011 with some seasons missing. We decided to examine just the full years of 2007 – 2010, since we expected to compare seasons when we began the analysis. We began with 1.5 million records.

A quick glance at the whole set of data, identified a strange outcome. There was a significant amount of negative consumption flow numbers, and some of them were quite large. We thought these might be just recording errors, but there were many negative readings. In questioning the PWA about this matter, they surprisingly wanted the negative readings to remain in the data set. They believed that the negative readings were corrections for previous incorrect charges. The negative consumption flows were caused by problems associated with reading the meters. Sometimes meters were incorrectly read or even read twice, resulting in overcharges to a customer. Rather than refunding money, the PWA wanted to reimburse the customers by recording negative consumption flow at a future time. Some of these negative readings were quite substantial, especially when they applied to a wholesaler, that is a company that is buying water from the Providence Water Authority for resale. The PWA was not always clear why negative readings would appear. They stated that mechanical devices can break, slow down, or in some usual circumstances run backwards, not to mention any human errors that might have occurred in data collection.

We could not agree with the PWA on this issue. We eliminated all records with a negative consumption flow, which would have been associated with some type of error, overcharging due to human error or machine error. We also decided to remove all records of a particular service location, which showed a negative consumption flow at any point in time. If a location had even one negative reading, then some type of error occurred at some point. We chose to eliminate all observations with that service location, even the ones with positive consumption flow. As it

turned out, about 94,000 records were eliminated. Those eliminated represented about 6% of the data. We still had over 1.4 million records in the data.

As with any data set, there are always concerns of outliers. Once all negative readings were removed, the other outliers would be readings that were very close to zero, representing zero water flow for an account or very large readings, representing wholesale accounts. These accounts would definitely skew the outcomes if they were not removed. We decided to focus our analysis on residential accounts and commercial accounts at this time. By eliminating the lower and upper first percentiles, we removed accounts with almost zero water flow and wholesale accounts, leaving well over 1 million records in the set.

INITIAL ANALYSIS

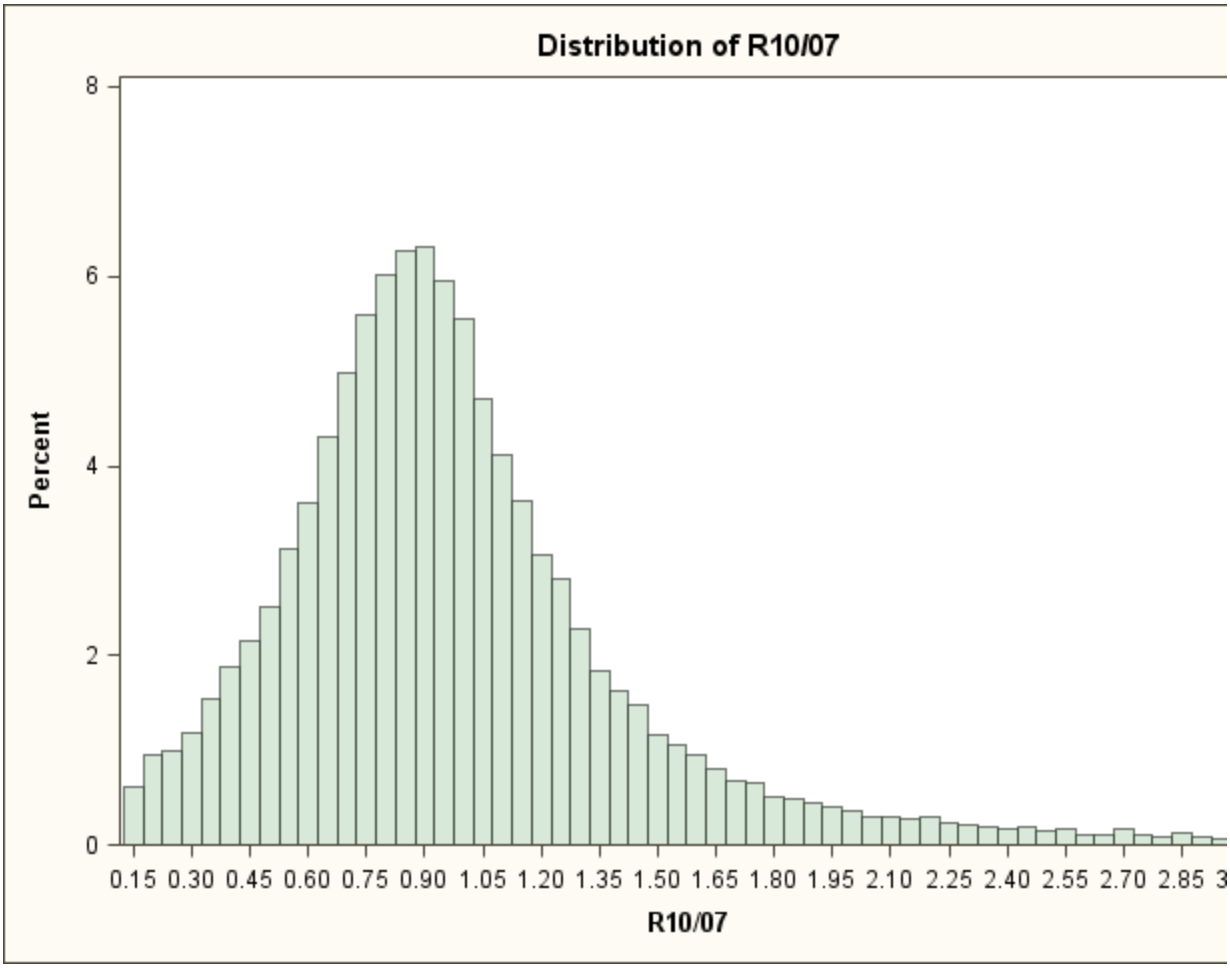
We began by choosing to compare consumption rates for the years 2007 to 2010 to get an overall sense of the data. We also chose to compare the rates for 2009 and 2010 to simply analyze the last full year of data. Thus two ratios were calculated for each service location. One ratio was the 2010 flow divided by the 2007 flow, the second was the 2010 flow divided by the 2009 flow. If the ratio for that service location was above one, for example 1.21336, then we had a 21% increase in that time span. If the ratio was below one, for example 0.90, then we had a 10% decrease in flow during the time span.

The following table gives the average of ratios for 2010 vs 2007, 2010 vs. 2009, and the average flows by location for 2010, 2009, and 2007, respectively.

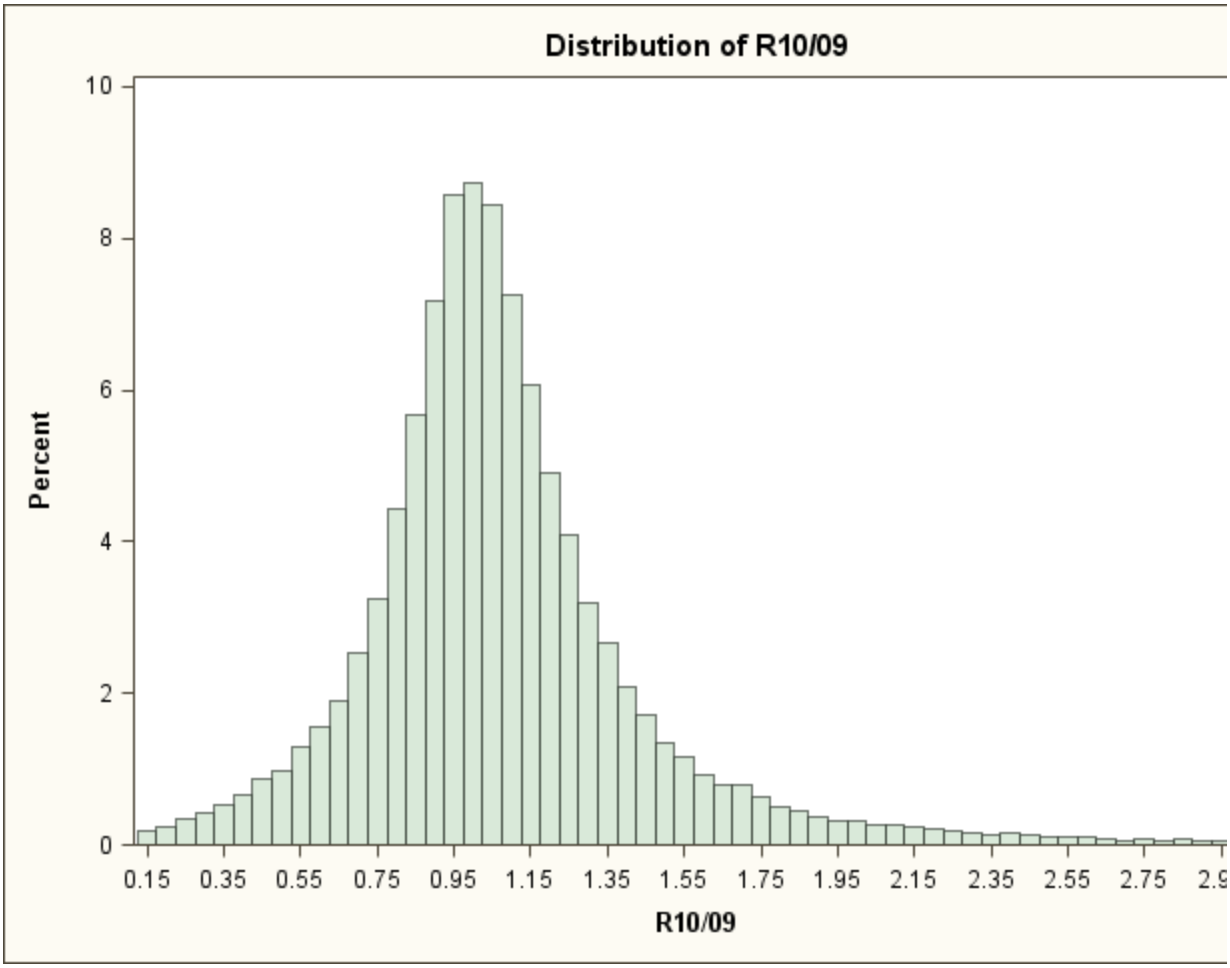
Variable	Mean	Std Dev	Minimum	Maximum	Sum	N	1st Pctl	Median	99th Pctl
R10/07	0.9731412	0.4436273	0.1301058	3.0000000	59673.02	61320	0.1966757	0.9087546	2.5434319
R10/09	1.0728776	0.3649223	0.1301938	3.0000000	65788.85	61320	0.3005172	1.0281208	2.4048630
SUM_of_flow	12708.20	11725.43	1.0000000	149930.00	779266791	61320	1061.00	9794.50	63081.00
SUM_of_flow1	12585.63	12115.18	1.0000000	149120.00	771751135	61320	1186.00	9459.50	65780.00
SUM_of_flow11	14567.74	14055.88	1.0000000	149890.00	893293748	61320	1419.00	10867.00	78620.00

Table of Average of Ratios for 2010 vs 2007, 2010 vs 2009, and average flows by location for 2010, 2009, and 2007

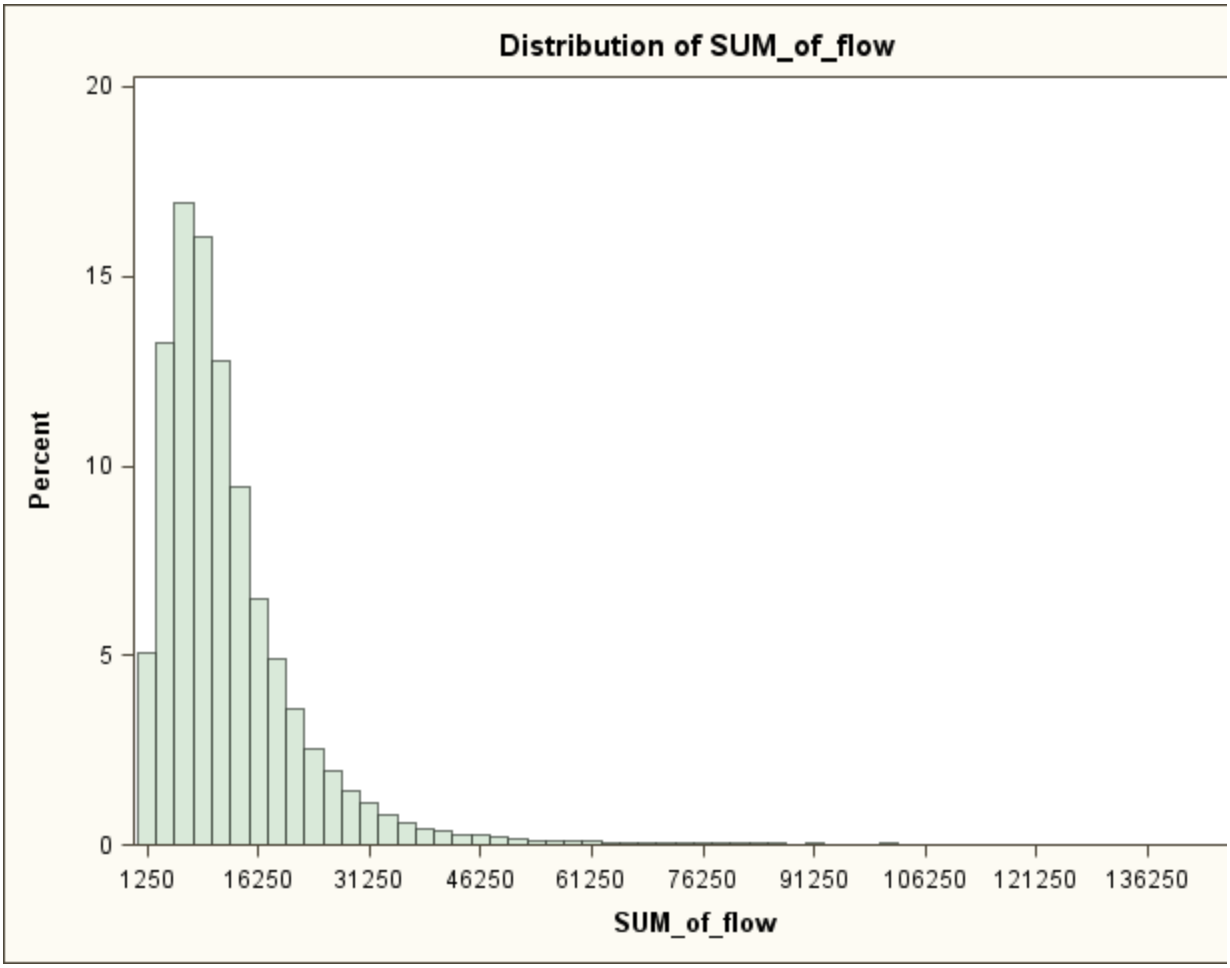
The following charts give histograms for individual ratios by location 2010 vs 2007, individual ratios by location 2010 vs 2009, and actual flow for all individual locations 2010, 2009, and 2007.



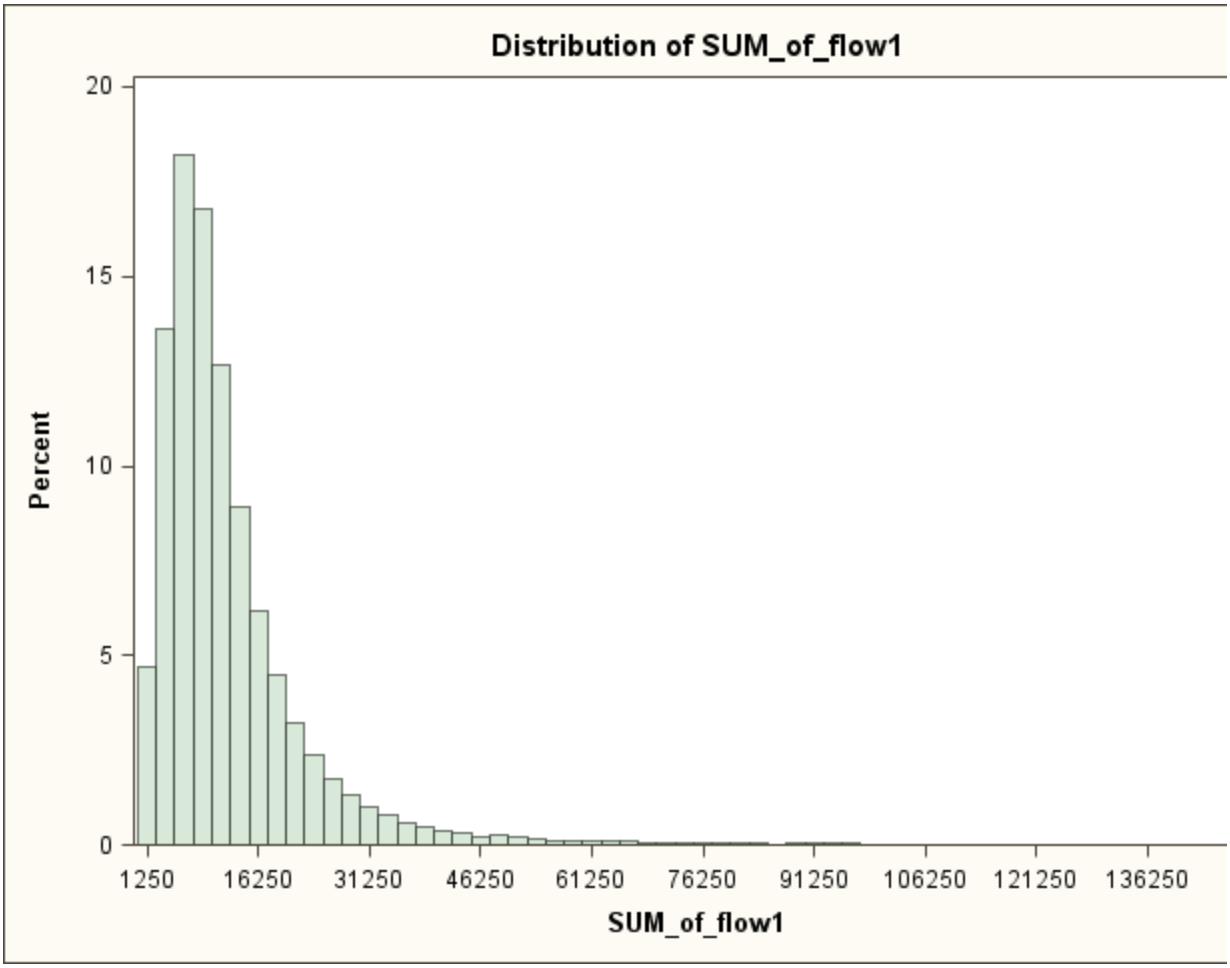
Histogram of Individual Ratios by Location for 2010 vs 2007



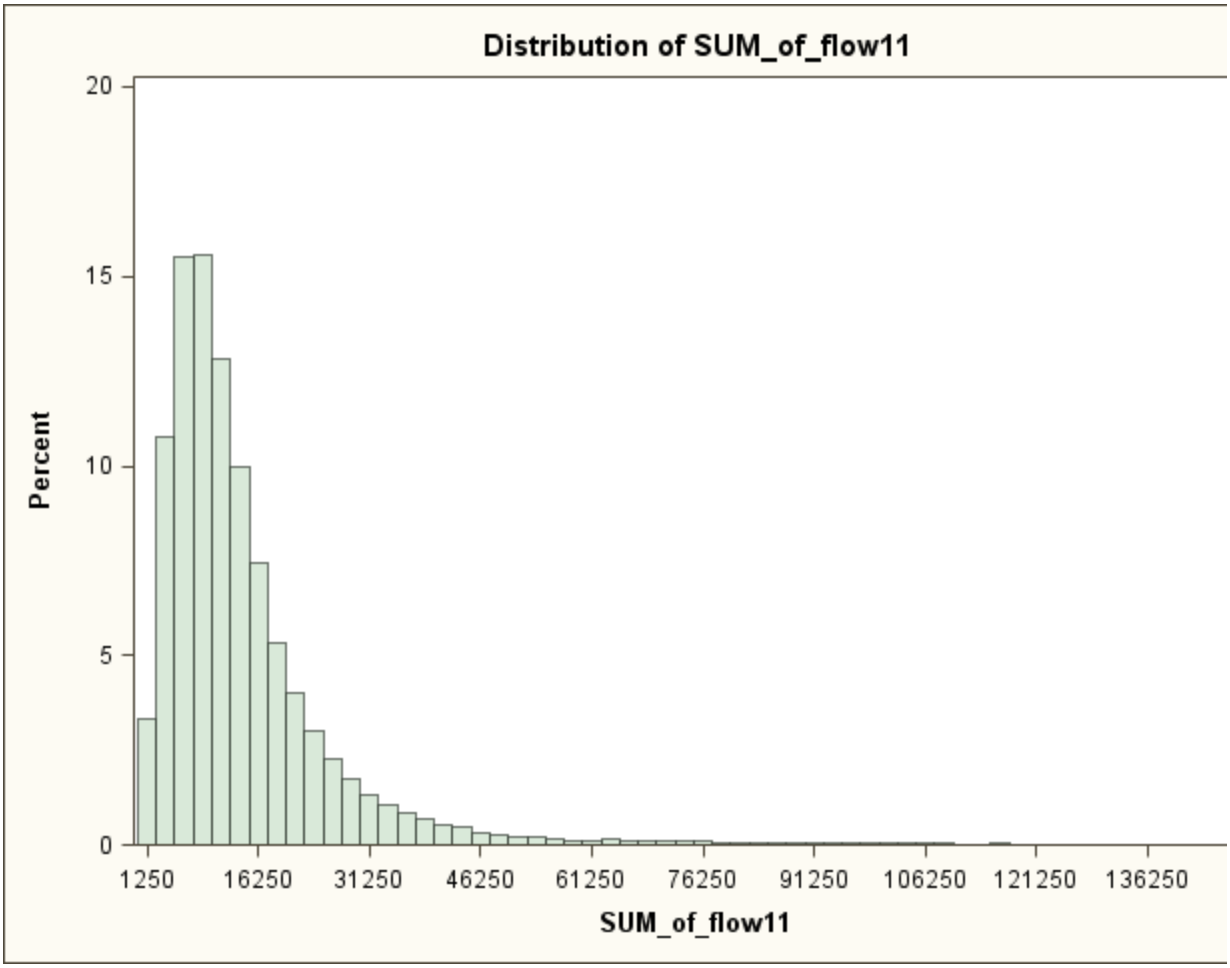
Histogram of Individual Ratios by Location for 2010 vs 2009



Histogram of Actual Flow for all Individual Locations - 2010



Histogram of Actual Flow for all Individual Locations - 2009



Histogram of Actual Flow for all Individual Locations - 2007

It should be noted that there was a lower consumption flow in 2009, which was a wet and cool year and a year of lower income. In 2007 comparison, the ratios of comparative flow were higher, and the economy was better that year.

CONCLUSIONS

The process of cleansing the data can take quite a long time. Analysts can expect that the raw data which they receive will have inherent problems. There is a continual back and forth conversation with the analyst and the client just to prepare the data. Furthermore, at the beginning of the project, a client often cannot formulate exact or particular questions of the data. The analyst must wrestle with the data to form clear cut questions that can actually be answered by analysis. We found that we needed to discover the problems with the data on our own. Each time we found a problem with the data, we would then speak with the PWA again. The PWA would often confirm what we had discovered, and they agreed that there was a problem with the data. Then no more information was given, until we discovered another problem.

It took many hours of building an understanding of the data content. Identifying what each variable means and what it does not mean is a long process. Some of the original information given to us was actually unreadable. For example, the problem with the service location and the account number was difficult to understand at the beginning. It turned out that the service location was of more importance to our questions than the account number.

The original question that the PWA wanted answered was whether or not a service location, residential or commercial, could be identified as a location that practiced conservative consumption. This was virtually impossible, and the question was unanswerable. There are too many variables that can force a consumption rate to increase or decrease. Children can move out of a household, a house can be sold to a larger family with more family members, a company can go out of business. These are all valid reasons why consumption flow can change, and it does not reflect whether people are using good consumption practices or not.

At the final meeting with the PWA, we could discuss statistical goals that would be more attainable, goals that they could achieve with the data at hand. The PWA does believe that as utilities become more expensive, the general public will become more interested in reducing their consumption of utilities and, in particular, their own water usage. They would like to begin to develop, in the long run, a price break system similar to the tax code system such that customers would be charged different rates at different consumption break points. This is now doable after more than a year of cleansing the data, and it will be the subject of future analysis.

A continuation of this work will include designing a new pricing system and comparing it to the present system that is now in place. In comparing the two systems we will be able to identify whether either system is better for a residential customer or a commercial one.

REFERENCES

1. Real Problems, Real Interactions, Real Outcomes
   Richard Tweedie and Sue Taylor
   Statistical Science , Vol. 13, No. 1 (Feb., 1998), pp. 1-3
   Published by: Institute of Mathematical Statistics
   Article Stable URL: http://0-www.jstor.org.helin.uri.edu/stable/2676708

2. Nonstatistical Aspects of Statistical Consulting
   Douglas A. Zahn and Daniel J. Isenberg
   The American Statistician , Vol. 37, No. 4, Part 1 (Nov., 1983), pp. 297-302
   Published by: American Statistical Association
   Article Stable URL: http://0-www.jstor.org.helin.uri.edu/stable/2682767

3. Statistical Consulting in a University: Dealing with People and Other Challenges
   Roger E. Kirk
   The American Statistician , Vol. 45, No. 1 (Feb., 1991), pp. 28-34
   Published by: American Statistical Association
   Article Stable URL: http://0-www.jstor.org.helin.uri.edu/stable/2685235

4. Data Cleaning: Problems and Current Approaches
   Erhard Rahm and Hong Hai Do
   University of Leipzig, Germany
   http://dbs.uni-leipzig.de
   Data Engineering
   December 2000 Vol. 23 No. 4 IEEE Computer Society.

5. Enterprise Guide (version 4.2), SAS Corporation, Cary, NC 2011.

6. http://www.provwater.com. Retrieved on August 22, 2011.

CAPITAL GAINS TAX RATES AND ENTREPRENEURSHIP

Vasanthakumar N.Bhat, Lubin School of Business, Pace University, New York, 10038
(vbhat@pace.edu)

This paper examines the impact of capital gains tax reduction on entrepreneurial activity. Both univariate and multivariate analysis indicate that a reduction in capital gains tax rate does increase likelihood of an individual starting a business.

Keywords – Capital gain tax rates, entrepreneurship, GEM survey

A capital gain represents the amount received from selling an asset such as a home, farm, ranch, or family business, over the price paid for it. The tax rates on capital gains have been a subject of considerable debate in the United States because of their impact on entrepreneurship and job

Table 1: Univariate Analysis of Entrepreneurial Activity Rates, 2002 and 2004

		% You are, alone or with others, currently trying to start a new business, including any self-employment or selling any goods or services to others		% You are, alone or with others, currently trying to start a new business or a new venture for your employer as part of your normal work		you have, in past three years, personally provided funds for a new business started by someone else excluding any purchases of stockes or mutual funds		You are, alone or with others, expecting to start a new business, including any type of self-employment, within the next three years	
		2002	**2004**	**2002**	**2004**	**2002**	**2004**	**2002**	**2004**
Age	18-24	**14.2**	10.2	9.5	**10.2**	5.8	**6.5**	**20.3**	13.1
	25-34	12.9	**23.3**	7.8	**12.7**	**4.9**	3.6	17.8	**23.2**
	35-44	**13.9**	11.4	6.6	**7.5**	**5.2**	3	**16.8**	14.1
	45-54	10	**11.9**	**6.3**	5.3	**4.5**	3.7	**12.6**	7.8
	55-64	8.9	**9.9**	3	**3.4**	4.5	**5.7**	8.6	**9.1**
	over 65	2.7	**3.7**	0.8	**1.5**	**4.4**	2.8	1.5	**2.4**
Gender	Male	12.4	**12.9**	8.2	**9.1**	**5.3**	4.9	**15.7**	13.8
	Female	8.6	**11.4**	3.3	**4.9**	**3.8**	3.1	**10.4**	10
Work Status	F-T, P-T	12.5	**14.9**	9.7	**9.9**	**5.4**	4.5	**15.6**	14.6
	Not Working	**11.2**	10.8	**3.1**	1.4	2.7	2.7	**14.6**	12.2
	Retired Students	3.3	**5.6**	**9**	8.1	2.8	**3.3**	3.2	**5.1**
Total		10.5	**12***	5.7	**6.9**	**4.5**	4	**13.1**	11.8

*Only total tested for statistical significance. Statistically significant at the 0.05 level.

Table 1: Univariate Analysis of Entrepreneurial Activity Rates, 2002 and 2004 (contd)

		In the next six months there will be good opportunities for starting a business in the area where you live		You have, in the past 12 months, sold, shut down, discontinued or quit a business you owned and managed, any form of self-employment, or selling goods or services to anyone	
		2002	**2004**	**2002**	**2004**
Age	18-24	**35.1**	28.6	3.2	**1.1**
	25-34	40.1	**43.3**	**3.1**	3.6
	35-44	40.4	**42.1**	4.3	**1.3**
	45-54	**36.2**	35.8	**2.6**	4
	55-64	29.7	**36.7**	**3.7**	8
	over 65	**20.6**	15.7	**1.2**	1.8
Gender	Male	39.5	**42**	3.5	**2.8**
	Female	**28.8**	27.6	2.5	**1.6**
Work Status	F-T, P-T				
	Not Working				
	Retired Students				
Total		34.3	34.6	3	**2.2**

creation. The purpose of this paper is to examine the impact of reduction in federal capital gains taxes in 2003 (also known as Bush tax cut and “The Jobs and Growth Tax Relief Reconciliation Act of 2003” or "JGTRRA" or Public Law No. 108-27, 117 Stat.752) on entrepreneurial rates in the U.S. This statute reduced the capital gains tax from rates of 8%, 10%, and 20% to 5% and 15%. We analyze the impact of these reductions using the *Global Entrepreneurship Monitor* (GEM) database. This database, unlike other databases that deal with firm level data, focuses on individuals and their involvement in entrepreneurial activities. This database is comprehensive as it is based on surveys of random samples of individuals on all phases of entrepreneurial activities, both formal and informal.

There is a considerable amount of literature on this topic. Gentry (2010) summarizes the impact of capital gains taxes on entrepreneurship. Most studies focus on aggregate impact rather an individuals decision to start new business.

Since the reductions in capital gains taxes were passed by the United States Congress by enacting Jobs and Growth Tax Relief Reconciliation Act of 2003 on May 23, 2003 and signed

into law by President George W. Bush on May 28, 2003, we ignore the survey results for 2003.So we evaluate the impact of the reductions by comparing survey results for 2002 and 2004. We compare the entrepreneurial activities based on the response to the questions relating to the starting of new businesses, providing funds for new business, expectation of starting of

Table 2: "You are, alone or with others, currently trying to start a new business, including any self-employment or selling any goods or services to other" (yes=1, no=0) as dependent variable

	Coefficient	Wald	Sig.	Odds Ratio
Male=1	0.199	7.237	**0.007**	1.22
Age category		31.463	**0**	
18-24	0.979	20.878	**0**	2.663
25-34	1.077	26.069	**0**	2.935
35-44	1.031	24.048	**0**	2.805
45-54	0.775	13.278	**0**	2.171
55-64	0.821	15.189	**0**	2.273
Knows someone personally who started a business in the past 2 years (yes=1)	0.995	180.936	**0**	2.704
Work Status		11.759	**0.003**	
Work:F-T, P-T	0.529	11.045	**0.001**	1.697
Not working	0.571	10.103	**0.001**	1.769
Education		2.167	0.538	
SOME SECONDARY	0.057	0.137	0.711	1.059
SECONDARY DEGREE	0.162	1.887	0.17	1.176
POST SECONDARY	0.125	1.262	0.261	1.134
Constant	-4.137	428.32	**0**	0.016
Year 2004=1	**0.686**	**60.225**	**0**	**1.986**
-2 loglikelihood	5277			
Chi-square	550		**0**	
# of observations	7882			

new businesses, and closing of businesses. We provide results based on univariate analysis and logistic regression analysis.

In Table 1, we provide comparisons of percent of people of age 18 and over who answered in affirmative to the questions relating to starting, starting for the employer, funding, expecting to start, seeing opportunities, and closing of businesses. We tested statistical significance of only

total percentages and we find that differences are statistically significant only for the percent of people who expect to start a new business.

Since Table 1 deals with only univariate analysis which does not take into account the impact of other variables, we perform logistic regression analysis to evaluate the impact of several variables simultaneously.

The logistic regression analysis indicates that the dummy for year 2004 is positive and statistically significant at the 0.05 level and log of odds ratio is close to 2. This means that it is twice more likely that an individual would start a business in 2004 compared to 2002 with everything else remaining constant.

Table 3: You are, alone or with others, currently trying to start a new business or a new venture for your employer as part of your normal work (yes=1, no=0) as dependent variable.

Variable	Coefficient	Wald	Sig.	Odds Ratio
Male=1	.658	42.065	**.000**	1.930
Age category		29.559	**.000**	
18-24	1.317	15.913	**.000**	3.731
25-34	1.172	12.612	**.000**	3.229
35-44	.968	8.533	**.003**	2.634
45-54	.944	8.061	**.005**	2.569
55-64	.471	1.817	.178	1.601
Knows someone personally who started a business in the past 2 years (yes=1)	1.003	104.616	**.000**	2.726
Work Status		40.305	**.000**	
Work:F-T, P-T	1.121	19.399	**.000**	3.069
Not working	.202	.441	.507	1.224
Education		22.537	**.000**	
SOME SECONDARY	.871	19.091	**.000**	2.390
SECONDARY DEGREE	.613	13.375	**.000**	1.845
POST SECONDARY	.398	6.022	**.014**	1.488
Constant	-5.987	336.693	**.000**	.003
Year 2004=1	.694	37.352	**.000**	**2.001**
-2 loglikelihood	3356			
Chi-square	442		**0.000**	
# of observations	7879			

In Table 3, we examine the answer (yes=1 and n0=0) whether an individual is starting a business for an employer. Here again, the dummy for the year 2004 is positive and the log of odds ratio indicates that the odds that an individual start a business for the employer is twice more likely in 2004 compared to 2002 all other things being equal.

In Table 4, we examine the answer (yes=1 and no=0) whether an individual alone or with others, expecting to start a new business, including any type of self-employment, within the next three years. The logistic regression output indicates that the dummy for 2004 is positive and statistically significant. In addition, the log of odds ratio indicates that the odds of an individual is expecting to start a new business in next 3 years is 1.5 times more likely in 2004 as compared to 2002 everything else being equal.

Unlike in other tables where we examined the starting of businesses, in Table 5, we examine the likelihood of businesses being closed. Based on the log of odds ratio, It is interesting to find that people of over 35 are 2 to 2.5 more likely to close their business as compared to below 35 everything else being equal. It is also interesting that if a person who knows that someone who has started business in preceding 2 years is 4 times more likely to close a business as compared

Table 4: You are, alone or with others, expecting to start a new business, including any type of self-employment, within the next three years. (yes=1, no=0) as dependent variable.

Variable	Coefficient	Wald	Sig.	Odds Ratio
Male=1	.330	21.699	**.000**	1.391
Age category		87.885	**.000**	
18-24	1.999	64.578	**.000**	7.383
25-34	1.927	60.514	**.000**	6.870
35-44	1.837	55.198	**.000**	6.280
45-54	1.477	35.076	**.000**	4.379
55-64	1.359	29.888	**.000**	3.893
Knows someone personally who started a business in the past 2 years (yes=1)	1.258	314.679	**.000**	3.518
Work Status		8.297	**.016**	
Work:F-T, P-T	.346	4.740	**.029**	1.413
Not working	.506	8.184	**.004**	1.659
Education		5.188	.159	
SOME SECONDARY	-.002	.000	.989	.998
SECONDARY DEGREE	.035	.090	.764	1.035
POST SECONDARY	.178	2.761	.097	1.195
Constant	-4.707	395.495	**.000**	.009
Year 2004=1	**.434**	**23.332**	**.000**	**1.543**
-2 loglikelihood	5552			
Chi-square	753		0.000	
# of observations	7818			

Table 5: You have, in the past 12 months, sold, shut down, discontinued or quit a business you owned and managed, any form of self-employment, or selling goods or services to anyone. (yes=1, no=0) as dependent variable.

Variable	**Coefficient**	**Wald**	**Sig.**	**Odds Ratio**
Male=1	.284	4.273	.039	1.328
Age category		12.378	.030	
18-24	.378	1.139	.286	1.459
25-34	.656	3.754	.053	1.927
35-44	.928	7.882	.005	2.529
45-54	.687	4.254	.039	1.988
55-64	.804	6.533	.011	2.234
Knows someone personally who started a business in the past 2 years (yes=1)	1.362	89.533	.000	3.903
Work Status		10.573	.005	
Work:F-T, P-T	-.441	3.269	.071	.643
Not working	.106	.147	.701	1.112
Education		6.850	.077	

SOME SECONDARY	.565	3.467	.063	1.760
SECONDARY DEGREE	.540	4.632	.031	1.715
POST SECONDARY	.620	6.798	.009	1.858
Constant	-5.197	243.853	.000	.006
Year 2004=1	.154	.739	.390	1.166
-2 loglikelihood	2017			
Chi-square	136		0.000	
# of observations	7900			

to person who does not know someone who has started a business in the previous 2 years. Work status is found to have no influence on the decision to close business. The 2004 year dummy is statistically insignificant indicating that business discontinuation rates are not affected by reductions in capital gains taxes.

Table 6: In the next six months there will be good opportunities for starting a business in the area where you live. . (yes=1, no=0) as dependent variable.

Variable	Coefficient	Wald	Sig.	Odds Ratio
Male=1	.371	46.916	.000	1.450
Age category		16.107	.007	
18-24	.190	2.086	.149	1.209
25-34	.336	7.003	.008	1.400
35-44	.438	12.213	.000	1.549
45-54	.278	4.984	.026	1.321
55-64	.195	2.599	.107	1.216
Knows someone personally who started a business in the past 2 years (yes=1)	1.144	436.663	.000	3.138
Work Status		15.116	.001	
Work:F-T, P-T	.387	13.753	.000	1.472
Not working	.246	4.027	.045	1.279
Education		30.858	.000	
SOME SECONDARY	-.454	15.674	.000	.635
SECONDARY DEGREE	-.335	14.708	.000	.715
POST SECONDARY	-.082	1.006	.316	.921
Constant	-1.667	224.012	.000	.189
Year 2004=1	.016	.047	.828	**1.016**
-2 loglikelihood	8297			
Chi-square	803		0.000	
# of observations	7037			

In Table 6, we examine perceived opportunities. In the logistic regression analysis, our dependent variable is yes=1 or no=0 to the question whether in the next six months there will be good opportunities for starting a business in the area where you live. Year dummy for 2004 is not

statistically significant indicating that capital gains tax cut does not have any impact on the perceived opportunities.

In conclusion, we should say that capital gains tax cut has positive influence on the starting of businesses. However, in terms of perceived opportunities, we do not find any positive impact. Because of terrorist attack in 2001, the economy in 2002 was not as good as 2004. Therefore, all improvements in entrepreneurial activities in 2004 may not be just as a result of reduction in capital gains taxes. However, comparison of entrepreneurial activities in 2002 and 2004, we do find statistically significant improvements in 2004 and we surmise that it could be because of reduction of in capital gains or policies enacted between 2002 and 2004 that encourage entrepreneurship.

William M. Gentry (2010) Capital Gains Taxation and Entrepreneurship, Williams College January.

GLOBAL REVIEW OF PARKING MANAGEMENT SYSTEMS & STRATEGIES

Pranav Joshi
Department of Operations & Information Systems
University of Massachusetts Lowell - Robert J. Manning School of Business
One University Avenue
Lowell, MA 01854
978-761-9813 | pranav_joshi@student.uml.edu

M. Riaz Khan
Department of Operations & Information Systems
University of Massachusetts Lowell - Robert J. Manning School of Business
One University Avenue
Lowell, MA 01854
978-934-2761 | m_Khan@uml.edu

Luvai Motiwalla
Department of Operations & Information Systems
University of Massachusetts Lowell - Robert J. Manning School of Business
One University Avenue
Lowell, MA 01854
978-934-2754 | luvai_motiwalla@uml.edu

ABSTRACT

Parking challenges have been a part of our society for a long time and traditional parking management strategies have come a long way. The problems associated with parking are common to most of us. This paper attempts to review globally implemented parking management strategies that leverage innovative technologies. Further, by examining a variety of parking management solutions from around the world, this paper aims to examine the shift in focus of modern parking management strategies. These solutions address conventional parking challenges local to their region, in an unconventional way. Finally, the paper performs a comparative analysis between traditional and modern (innovative & technology-driven) parking management strategies/systems. The motivation for this research paper is to identify the positive commonalities of the innovative approaches, which would aid in designing future parking management models.

Keywords: Global Parking Management, Review of Parking Strategies

INTRODUCTION

The Parking Problem

Everyone who owns a vehicle needs a place to park. As the number of vehicles grows, so does the need for parking. According to the U.S Department of Transportation's Federal Highway Association (USDOT FHWA), the number of publicly owned motor vehicles has risen, more than 20% from 3,493,570 in 1994 to 4,224,542 in 2009, as shown in Figure 1. Vehicle ownership rates in densely populated countries are even higher. The conventional parking management strategy has been to increase the existing capacity. Other innovative strategies address the rising vehicle ownership trends more effectively. Due to this growth, parking space has become an integrated component of the building and road infrastructure. A lack of adequate, easy-to-find, and inexpensive parking facility aggravates operators who must routinely travel to carry out a gamut of their daily activities. Ignoring the issue or under-investing in parking management is likely to contribute to increase in traffic congestions, traffic violations, vehicular accidents and injuries, and waste of time and money.

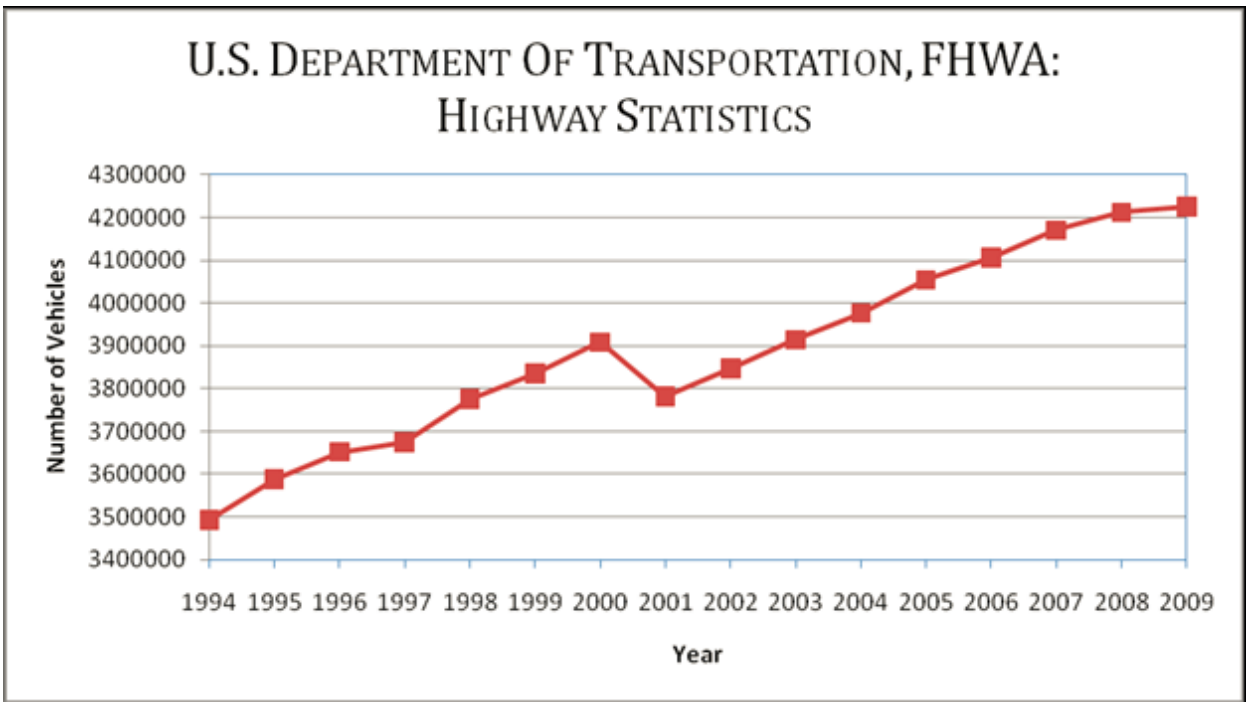


Figure 1

Lack of effective parking management creates problems for everyone. Either improvements in the existing solutions require significant financial commitments or they are difficult to implement due to their complex nature. Individual entities such as cities, firms, stadiums, and subways have their own particular parking arrangements. However, there is no standard parking management skeleton. Each setting is different and what works for one organization may not work for another. Two motives are common among all who wish to provide a parking solution;

first, they all strive to address their unique parking needs and second, earn the most from investment in their parking facility. Due to the substantial investment and complexity involved in parking management, on the other hand, many institutions choose to outsource. These institutions are either unable or unwilling to deal with the challenges of operating and maintaining a parking facility.

The planning and resources needed for setting up a parking facility vary significantly across organizations. It could be as simple as posting a person outside a gated lot, or it could be as complex as managing the vehicles of all commuters at an airport. The amount of resources, including parking space is, of course, limited. Instead of searching for individual parking solutions, energies should be expensed on having a comprehensive view of the problem and identifying a standard solution that would be applicable to most parking systems.

Goals/Scope of this paper

This paper identifies the common strategies emerging from parking problems in various countries. While, peculiar parking challenges are dealt with by using unconventional methods, most conventional strategies used today attempt to solve the problem by relying on traditional approaches. Such methods may include expanding parking capacity, investing in parking infrastructure, providing incentives to use public transportation, and penalizing those who violate the parking rules. This paper attempts to review, discuss, and perform a comparative analysis of parking management strategies used globally. Further, it evaluates and compares the innovative solutions used to address the parking problems and traditional approaches. Next, the paper highlights some common strategies and features that emerge from the solutions. Lastly, the paper analyzes the strengths and weaknesses of the solutions utilized in these strategies and suggest few recommendations.

PARKING MANAGEMENT

Parking systems routinely experience parking related challenges, especially in the urban and metropolitan areas. The most significant of them is the availability of space. Either the operators cannot find parking slots or the capacity is under-utilized. Many institutions use management tools to solve their particular parking issues. According to some transportation authorities, such as Victoria Transportation Policy Institute (VTPI) and Texas Transportation Institute (TTI), parking management refers to policies and programs that result in efficient use of parking resources (VTPI) or strategies aimed at making better use of available parking supply (TTI). These definitions are very broad. However, for a parking management strategy to be effective in solving parking problems, the following must be true:

i. A comprehensive needs analysis of the facility must be conducted
ii. A small scale prototype be deployed and the findings carefully analyzed

There are several types of parking management systems in existence. Discussed below are some the traditional parking management strategies.

Overview of Traditional Parking Management Approaches

Effective parking strategy relies on either a specific type that meets the needs of a system or a combination thereof. Reviewed below are some of the currently practiced parking management approaches:

1. *Brick and Mortar Facility:* Majority of institutions use these traditional parking facilities either investing in large spaces of land or constructing large buildings to accommodate vehicles. This approach has been the tried and true approach. It addresses several parking management challenges effectively.
2. *Preferential Parking:* Use of this approach is to make the most of a given parking space. A good example would The Los Angeles Department of Transportation. They implement and enforce Preferential Parking Districts in residential areas where employees and customers of nearby businesses and attractions regularly park for long periods of times making it difficult for residents to find parking. Preferential Parking Districts restrict parking for all motorists, but area residents and their guests are exempt from the special parking restrictions if they purchase and display Preferential Parking Permits [1]. Another example would include spaces reserved for people with two or more individuals in a vehicle, similar to a HOV car pool lane.
3. *Price Discounts:* This approach usually applies to carpools and or short-term parkers. This is desirable for change-mode parking facilities [2] such as airports and subways, where commuters often park their vehicles before using public transportation services. In addition, carpool lanes on busy highways offer timesaving.
4. *Change-mode Parking Facility:* This approach is perfect for institution such as airports, subways, park-and-ride shuttle services. All of these institutions require the incumbents to locate a parking space promptly and continue commuting in a public transportation vehicle [2].
5. *Disincentives:* On those who contribute to traffic-congestions (illegal parking, parking near an expired meter), the state and local governments may impose citations. For example, a person is fined for travelling alone in a car pool lane that requires two or more passengers
6. *Satellite Parking:* This approach is also known as remote parking or off-site parking. This approach is used when the place to park is far away. Usually in this arrangement, a parking facility is shared with other entities or there are other incentives involved. Logan airport in Boston, for example, provides shuttle services to neighboring cities and towns. Another popular example is a free shuttle service to any casino is Las Vegas.
7. *Shared Parking:* This is a popular approach to address a need for large space requirements with significantly reduced investment. Typical examples include universities, business parks, and privately owned doctor's practices.

When parking management strategies are appropriately applied, they can effectively manage the number of parking spaces required in a particular situation and they provide a variety of undeniable economic, social, and environmental benefits. Each of these strategies, however, has its own strengths and weaknesses to be discussed later.

Challenges of Traditional Parking Management Approaches

It is well known that all of the above-mentioned traditional approaches are used to address different types of parking needs. Furthermore, all of these approaches come with their set of challenges as shown in Table 1 below:

Table 1

	Cost to Institute	Cost to Consumer	Time to Locate Parking	Staffing	Popularity
Brick & Mortar	Very High	Low	Moderate	High	Moderate
Change-mode	Moderate	Low	Moderate	High	Low
Preferential	Very-High	High	Low	High	High
Price Discounts	Low	Low	Moderate	Low	Moderate
Disincentives	Low	Moderate	NA	Moderate	Low
Satellite	High	Moderate	Moderate	High	Moderate
Shared	Moderate	Moderate	Moderate	Moderate	High

INNOVATIVE PARKING MANAGEMENT APPROACHES (METHODS)

FIGURE 2

Traditional parking management approaches address the fundamental parking space problem by expanding the parking capacity. This solution works however, is not efficient. Example, a university typically has most of the classes during the daytime. Most of the students park their cars and occupy all or most of the available spaces. However, in the afternoons or evenings those same spaces are vacant. In other words, the parking capacity is underutilized. Innovative parking management approaches have a different focus. One of the goals of these modern day parking management strategies has been to increase the utilization factor (usage/occupancy rate) of the existing capacity, Figure 2. They accomplish this by many ways:

1. Delivering current parking information to its users efficiently
2. Identifying peak/busy times and determine methods to maintain constant occupancy (renting, sharing, & leasing available parking spaces)
3. Implementing demand responsive pricing methodologies (example, SFPark)
4. Analyzing past parking trends and using that data to predict future parking availability (analytics)

Sixteen parking management strategies from different nations are reviewed. These strategies are divided in three groups:

Policy Based Solutions

FIGURE 3

These are parking management solutions governed by institution-determined policies. FutureLink, the maker of **Parking Guidance Systems** is a Dubai based company that supplies to the major cities in the UAE. It manufactures a sensor supported parking system, as shown in Figure 3. This system provides current parking availability update on screens placed throughout the parking facility. This real-time visual presentation of free spaces allows for easy parking and increases the utilization factor for a parking facility. In addition to the time savings, the system offers direct guidance, as depicted in Figure 4, to the destination via electronically controlled variable message signs. Another parking management solution from FutureLink, **Mobile Car Parks** caters to short term parking problems, especially when there is limited construction time and parking capacity. This system constructs parking lot buildings with the mindset that they are temporary. While these buildings, shown in Figure 5, are weather resistant and sturdy, they do not need the typical construction cost of a permanent parking structure. Reuse of buildings is possible and is easy to add capacity. Omnitec is another Dubai based company, which provides **Automated Number Plate Recognition** System & Access Control Management. This is extremely useful to institutions to enforce policy based parking management. This solution relies on a camera mounted at the parking gate, as shown in Figure 6. The camera scans the vehicles number plates, checks the internal database for a matching number, and controls access to the facility.

FIGURE 4

FIGURE 5

FIGURE 6

Another solution from Omnitec is the **Signal Ramp SP-143 Detector** [3] shown in Figure 7. This device is mounted throughout the parking facility. It constantly monitors the parking spaces for availability and then relays that information to the parking management system. This system then provides parking information in real-time.

FIGURE 7

In Ireland, there is heavy reliance on industrial size parking management solutions. These solutions often require large startup costs and on-going maintenance. Easy Park Parking Solutions provides some unusual parking management solutions discussed below, to use the limited parking spaces and strict institution policies. **Parking Boy Model** [4] is a parking space protector designed for disabled, VIP, and designated spaces. The goal here is to make it impossible for an unauthorized vehicle to park in a reserved space. The uniqueness lies in the radio-controlled hoop-shaped bar for individual parking spaces. When a car approaches a parking space, a hand-held radio transmitter lowers the device, as shown in Figure 8. When the car leaves the spot, the device acts like a barricade and prevents unauthorized entry. These are some examples of innovative parking management solutions, which provide policy based parking management addressing a variety of parking needs.

FIGURE 8

Technology Driven Solutions

These solutions provide parking management using advanced technological innovations. Known for their efficiency, innovativeness, and several economic benefits, in Australia, there are many technology-driven parking management solutions. One of the reasons is the size and uniformity of the highway infrastructure. In addition, a nation of 23 million, Australia has a literacy rate of 99.1%. This, coupled with the significant investments in research and development, makes Australian innovation more productive. Park Mobile International is a technology company that has made clever use of the Australian parking meter infrastructure. The street parking meters in Australia have a unique code. **ParkMobile** [5] was created as a Smartphone application that allowed the user to pay for parking via their phones. User was required to input the parking meter serial number and the amount of time. The application would then transmit that

FIGURE 9

information to the city police who would receive instructions not to impose citations to the parked vehicle. S & K Car Park Management Pty Ltd is the maker of **SecurePark** [6]**.** Similar to ParkMobile, the key difference is that SecurePark uses a mobile application to locate available parking across the city. It displays available parking on a map based on user's location, as shown in Figure 9. Its application capability to pinpoint the cheapest available parking in the vicinity is also an option. Similar to SecurePark is a parking management solution called **SFPark** [7]**,** developed in order to improve the city's parking facilities in San Francisco. This system also relays real-time parking information to the user's smartphones, as shown in Figure 10. However, what makes the system unique is the focus on increasing the utilization factor using technology. They call is demand-responsive pricing. The system responds to bring the supply and demand levels in balance by periodically adjusting meter and garage prices. Demand-responsive pricing encourages drivers to park in underused areas and garages, shifting demand from overused areas.

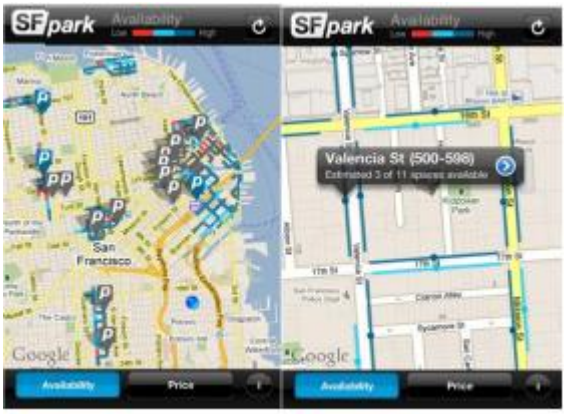


FIGURE 10

The abovementioned examples illustrate the use of technology to boost the utilization factor of existing parking capacity. In addition, they also provide efficiency, economic benefits (tangible/ intangible), timesaving, and safety to consumers. Technology driven parking management solutions like the ones seen above are new and are gaining popularity with generations after the baby boomers.

Economic Driven Solutions

These solutions are motivated by the parking needs and other external economic and environmental factors (such as laws limiting area used by parking spaces). Here the focus is not the use of technology but adherence of the external factors. As an example, the City of Lowell implements **AURA** [8], a smart parking meter developed by Metric Parking. The primary purpose of this modern parking meter is to prevent theft. The city had a problem of thieves stealing parking meters and therefore implemented this system, which is more secure than traditional coin operated meters. In addition, this meter manages multiple car spaces and accepts multiple modes of payment. Table 2 shows the key differences between traditional and AURA parking meters:

Table 2

Traditional Parking Meter	AURA Parking Meter
1. Price ranges from $362-$700 per meter	1. Cost of one meter is ~$10,000
2. Accepts one mode of payment (quarters)	2. Accepts multiple modes of payment
3. Not secure and may be stolen	3. Securely bolts on the ground with Anti-Theft device
4. One meter for one space	4. One meter for multiple spaces
5. Inconvenient collection of proceeds, individual meters need to be emptied	5. Armored car from bank collects proceeds from multiple space all at once
6. Cost of 20 spaces would range $7240 - $14000	6. Cost of 20 spaces would be ~$10,000, the cost of one unit

In Turkey, Katopark is a manufacturer of mechanical auto parking lift systems. This manufacturing company sells customized parking management solutions to malls, compact apartment complexes, office buildings, and other institutions where space is limited or restricted. However, these systems are expensive due to the need for customization, regular maintenance, and staff to help parking users. **Puzzle Parking** is a system manufactured by Katopark. This is a parking solution based on stacking the cars on floors on top of each other, as shown in Figure 11. Another solution from the same company is called **Square Parking.** This system is similar to Puzzle Parking however is customized to use the narrow basement spaces to its maximum capacity. The following figure (Figure 12) shows the system that incorporates car elevators.

FIGURE 11

FIGURE 12

FIGURE 13

A cost effective alternative to the above-mentioned

Puzzle Parking and Square parking systems is Easy Park Parking's **Easy Park VSS** that relies on stacking vehicles. As shown in Figure 13, this system is useful for doubling or even tripling the parking capacity.

Australian Parking and Revenue Control (APARC): This Company is a distributor for Parkeon, a company that manufactures high tech parking meters. One of its flagship parking meters is **Strada® NEOPS**. It is different from the AURA parking meter because APARC connects TicketManager™, a handheld enforcement solution to the vehicle sensors, which instantaneously provides the on-duty Parking Enforcement Officer (PEO) with the real-time occupancy status of each parking space. When a vehicle drives into a parking stall, vehicle sensors transmit this information to the back office. This is communicated in real-time to the handheld as a color-coded space list, which the PEO uses to identify the status of each space in a given area. The PEO is then able to quickly target and pinpoint the vehicles in violation [9].

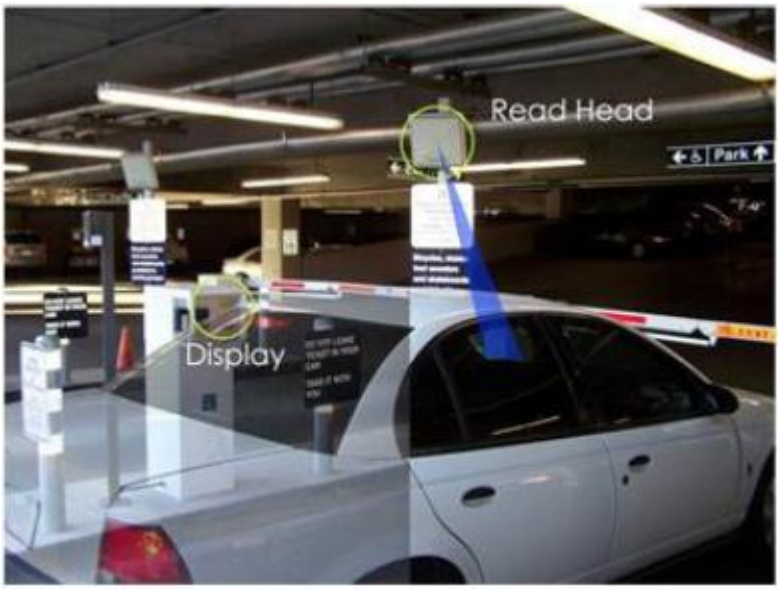


FIGURE 14

University of Arizona, a state university uses **Radio Frequency Identification (RFID) tags** across its campus parking lots. RFID antennas installed at entrance and exit, as shown in Figure 14, communicate by radio frequency with the tags attached to vehicles. The system sends the tag number to a computer database for verification. Upon successful verification, the system allows entry. University of Massachusetts Lowell uses the RFID based **U-Card**, which eliminates the need to install RFID receivers throughout the parking lots and inside vehicles. The RFID chip is present in the ID card issued to faculty, staff, and students. The ID card is scanned at the entrance and authorized vehicles are permitted inside the parking facility. The campus parking lots use automated entry and exit systems, shown in Figure 15. In addition, at every entrance and exit, there are underground devices (load loops) which detect the presence of a vehicle.

FIGURE 15

Parking management solutions are most economic-driven solutions. One of the primary factors is the nation's populations. Based on the data provided by the World Bank, the current population of China is 1.331 billion [10]. In China, there is a serious need for parking

management. The culture in China associates car ownership with high status in society. Therefore, more cars a family owns the higher the status it enjoys within the society. This has led to an explosion of number of vehicles in recent decades. According to a local newspaper, the number of cars in China has reached 100 million [11].

Today China is the most populated country in the world and, with the exponential growth, effective parking management is a keenly serious challenge. A good example of an economic driven parking management solution comes from a Balizhuang (suburb of Beijing) based company. They have introduced a parking space guidance system called **Nestca**. It was developed to curb the growing problem of a lack of local parking spaces. The web application allows users to view the large number of parking spaces left vacant by the local owners who drive out for work. The website [12] constantly displays available parking spaces as shown in Figure 16. People driving into the town use these vacant spaces. Nestca connects these commuters to the parking owners and provides them a digital map that highlights the rental parking slot's location in a user-friendly interface and allows searching a spot by locality or rent.

FIGURE 16

The review of the above-mentioned sixteen solutions demonstrated the additional focus on increasing the utilization rate of parking management solutions. These innovative parking solutions have addressed a wide spectrum of parking related challenges in their own unique ways. The differences in these solutions were present due to several factors such as country of origin, economic and environmental factors, policy factors, and technological factors.

Advantages/Disadvantages of Innovative Parking Management Approaches

As seen, parking needs vary across the spectrum. Every condominium complex, business organization, church, school, city, hospital, airport, and subway within a country, for example, has different parking needs. Therefore, to provide an effective parking management solution, there must be a needs analysis of that particular entity. However, all entities that seek effective management solutions could benefit from innovative parking management solutions, which focus on increasing the utilization factor. Table 3 provides a summary of the sixteen innovative parking management solutions discussed above:

Table 3

Detection Technology	Known Parking Installations	Advantages/ Disadvantages	Types of Facilities	Facility Scale	Maintenance Needs	Installation Costs
Parking Guidance Systems	Emirates Headquarters, Dubai International Airport, UAE	100% utilization rate, efficient, Good for policy-based solutions. Costly installation, Not Scalable,	Structure or Surface	Best for large buildings	High	High-varies
Mobile Car Parks	None	Large parking capacity, Cheaper than permanent parking facilities, Reusable, not scalable	Structure	Large Buildings	Moderate	High-varies
Automated Number Plate Recognition	The Pearl, Qatar, UAE	Highly secure, Good policy driven parking management	Structure	Small, medium, & large structures	Low - Moderate	Moderate
Signal Ramp Detector	Shangri-La Hotel, Dubai	Accurate Real-time Space availability, High resource consumption, 1 unit for 1 – 5 spaces	Structure	Large Structure	High	High
Parking Boy Model	Irish Parking Association	Ideal for reserving spaces, None to minimal staffing, Efficient. Costly, 1 unit for 1 space	Surface only	Small to Medium areas	High	High
ParkMobile	Most cities, Australia	Easy, inexpensive, efficient	Surface	Large Areas - Cities	Low	Low
SecurePark	Most cities, Australia	Inexpensive, efficient	Surface	Large Areas	Low	Low
SFPark	Prototype in San Francisco	Demand responsive parking, real-time info. Does not cater to non-Smartphone users	Surface	Large Areas - Cities	Moderate	Moderate
AURA Parking Meter	City of Lowell –downtown area	Covers Multiple spaces, Secure, Accepts all Payment types. Expensive to install, Hard to access during bad weather	Surface	Large Areas - Cities	Moderate	High
Puzzle Parking	None	Ideal for small areas	Structure	Small Areas/Tall Structures	High	High
Square Parking	None	Ideal for small areas	Structure	Small Areas/Tall Structures	High	High

[Table 3, continued]

[Table 3, continued]

Detection Technology	Known Parking Installations	Advantages/ Disadvantages	Types of Facilities	Facility Scale	Maintenance Needs	Installation Costs
Easy Park VSS	None	Ideal for small areas	Structure	Small Areas/Tall Structures	High	High
Strada® NEOPS	City of Boadilla near Madrid	Covers Multiple spaces, Secure, Accepts all Payment types. Expensive to install, Hard to access during bad weather	Surface	Large Areas	Moderate	High
RFID Device Tag	University of Arizona	Very Secure, Efficient entry/exit. Requires users to buy RFID device	Structure	Medium Areas	Moderate	High
Nestca	Major cities, China	Cost-effective, Helps everyone, increases utilization factor. Privacy concerns, no Smartphone application	Surface	Large Areas	Low	Low

Thus, as this review shows, there are several innovative parking management systems, which make good use of information technology (IT) and available resources.

Each of these solutions has certain strengths, which, of course, could be improved, and weaknesses that could be eliminated.

However, the most successful ones are those systems that focus on increasing the utilization factor of the existing parking capacity, such as Nestca, ParkMobile, SFPark, and SecurePark.

These solutions have made a clever use of today's IT and addressed their unique parking challenges.

In addition to the technologies summarized above (Table 3), Table 4 [13] below provides information on some other innovative technologies and their distinct features:

Table 4

Summary of Detection Technologies

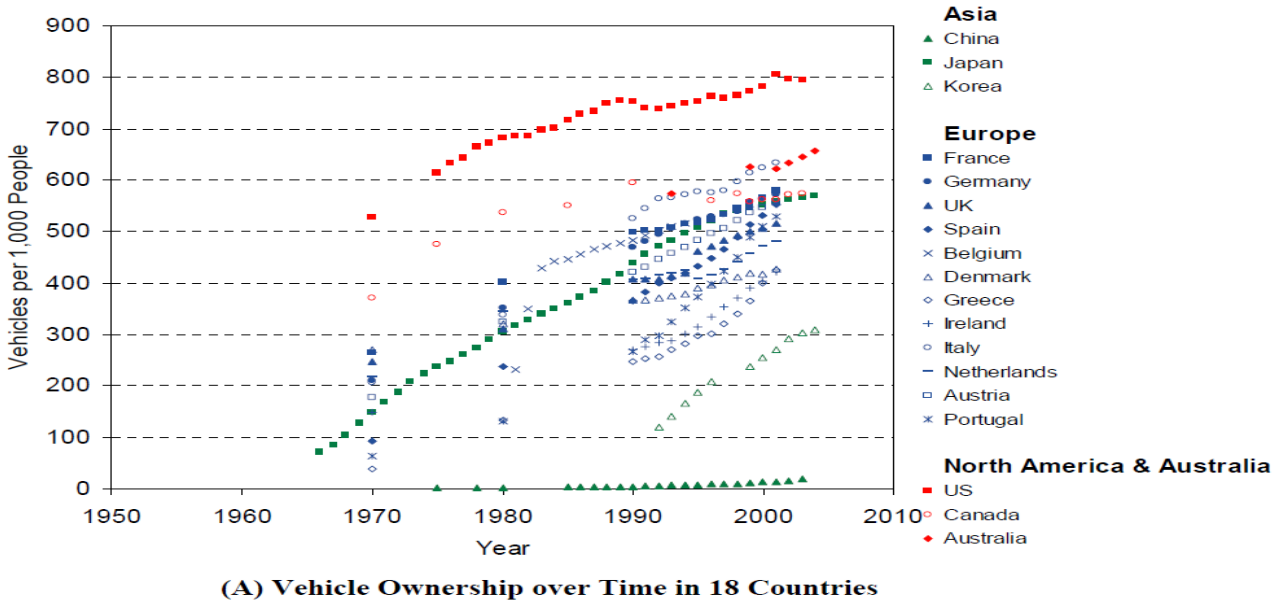
Detection Technology	Known Parking Installations	Advantages & Disadvantages	Types of Facilities	Facility Scale	Maintenance Needs	Installed Costs
Induction Loop	Metra (Chicago), MSP airport, Santa Monica, others	Reliable, widely deployed; Must be installed below pavement	Surface or structure	Best for portal detection	Low maintenance, can be damaged during excavation	$750 - $1,500
Magneto-meter	San Francisco, Portland airport	Inexpensive, can be installed in or on pavement, can operate on battery power with wireless connection. One detector needed per space; eventual need for battery or sensor replacement	Surface or structure	Scalable; typically one detector per space	Battery life approximately five years	$300
Ultrasonic	Baltimore airport, Munich airport, others	Flexible, can be installed over right-of-way Expensive; power requirements	Structure			$600 - $1,900
Passive Infrared	None	Effective for measuring speed Dependent on visibility conditions	Surface or structure			$700 - $1,200
Active Infrared	None	Can detect multiple zones; Dependent on visibility conditions Expensive	Surface or structure			$6,500
Video	Seattle-Tacoma Airport	Can use existing equipment. Expensive	Surface or structure	Best for small zones or portal detection		$5,000 per unit – up to 30 detectors per camera
Microwave radar	None	Can detect multiple zones, not weather sensitive. Expensive	Surface or structure			$800 - $3,300

GLOBAL REVIEW OF PARKING SOLUTIONS (EXAMPLES)

Comparative Analysis by Country

Table 5

	Australia	China	Ireland	Turkey	UAE	USA
Solution(s)	Secure Park, Park Mobile, Strada NEOPS	Nestca	Parking Boy Model, Easy Park VSS	Puzzle Parking, Square Parking	Parking Guidance Systems, ANPR, Signal Ramp SP-143 detector	AURA, SFPark, RFID Device tags, RFID ID cards
Approach	Technology Driven	Economic Driven	Policy-Based, Economic Driven	Economic Driven	Policy-Based	Technology, Economic Driven
Population [14] (World Bank, 2010)	21.8 million	1.33 billion	4.45 million	74.81 million	4.6 million	307 million
GDP [15] (World Bank, 2010)	$924.84 billion	$4.99 trillion	$227.19 billion	$614.6 billion	$230.25 billion	$14.12 trillion



(A) Vehicle Ownership over Time in 18 Countries

Figure 17 [16] (Wang, Huo, Johnson, & He, 2006)

Comparative Analysis of Traditional vs. Innovative Parking Management Solutions

Innovation and technology are becoming more important as natural resources such as land are dispensed. Increasing the utilization rate is more practical and efficient than increasing physical capacity. Most parking management solutions have not considered the utilization factor. Because of that, they are not using the available resources efficiently. In addition, this type of a setup has several drawbacks:

- Heavy reliance on capacity may lead to high initial and on-going investment costs
- Large capacity requires trained personnel for its maintenance
- Scarce availability during peak hours & underutilized capacity during non-peak hours
- Impact on the environment:
    - Building large parking facilities use precious land and other resources
    - Frustrated users waste fuel trying to locate a parking space

Difficulties Linked to Increasing Utilization Rate

So why do institutions neglect the utilization rate? In the past, parking management providers have ignored the utilization factor. This is because the solutions were designed keeping in mind the capacity. For example, the mall would design the parking lot taking into consideration the traffic during the busiest time of the year. This trend was due to the following reasons:

- **Traditional mindset:** Historically, institutions have addressed the space issue by adding more space. This has been the tried and true method of dealing with the issue. Institutions have devoted large number of resources to address the parking problem
- **Investment for redesign:** Some institutions are trying to increase the utilization factor. However, the resources needed of redesign of a parking solution are seldom too high. Examples are investments in modern technology, knowledgeable staff, etc.
- **Lack of technological knowhow:** Some institutions find it hard to increase their parking utilization rate. This is because they have a traditional mindset and or lack the technological knowhow

Emerging Common Features/Strategies from Review

Table 6

Parking Solution	Utilization Rate Focused	Customer-Centric	Easy to Deploy	Life Expectancy
1. Parking Guidance Systems	Yes	Yes	No	~5 years
2. Mobile Car Parks	Yes	Yes	No	5-8 years
3. Automated Number Plate Recognition	No	Yes	No	5-6 years
4. Signal Ramp Detector	Yes	Yes	No	3 years
5. Parking Boy Model	No	Yes	No	5-6 years
6. ParkMobile	Yes	Yes	Yes	>20 years
7. SecurePark	Yes	Yes	Yes	~5 years
8. SFPark	Yes	Yes	No	Data Unavailable
9. AURA meters	No	Yes	No	~12 years
10. Puzzle Parking	Yes	No	No	>20 years
11. Square Parking	Yes	No	No	>20 years
12. Easy Park VSS	Yes	No	No	>20 years
13. Strada NEOPS	No	Yes	No	~12 years
14. RFID Tags	Yes	Yes	No	>20 years
15. RFID Cards	No	Yes	Yes	>20 years
16. Nestca	Yes	Yes	Yes	>20 years

Table 6 above provides a representative list of sixteen parking management solutions. Innovative parking management strategies around the world either are one of the solutions mentioned above or are some variation of the same. Furthermore, the table shows how the solution providers have shifted their focus in terms of addressing the longstanding parking challenges. Out of all parking solutions reviewed in this paper:

- 68.7% of the solutions have focused on increasing the utilization rate
- 81.2% of these solutions are customer-centric
- 56.2% have a life expectancy of more than ten years

The data in Table 6 indicates that even though the parking management solutions are very different in their operation and implementation, they share some common strategic goals. These goals include but are not limited to:

- Boosting the utilization rate
- Being increasingly customer-centric
- Leveraging innovative technology to:
  - Reduce costs
  - Facilitate ease of use
  - Increase life expectancy

CONCLUSION/FUTURE DIRECTION

The versatility of the innovative parking solutions allows tremendous flexibility when implementing the various parking management solutions. There are a few ways the approach may be improved:

- Developing hardware, which is cost-effective, would be a major improvement. Due to the current high investment, costs associated with real-time parking related hardware such as sensors, RFID chips, etc. are high
- Reducing the costs associated with on-going maintenance would also help bring the cost down and therefore lead to wide acceptance
- Researching ways to increase the system uptime and study parking trends more efficiently will help make this model better

The innovative parking solutions approach has a wide range of applications. This approach may be applied in many types of scenarios such as university parking lots, street parking, airport parking, commercial car parks, etc. Of course, minor alterations to the required hardware are needed but the core functionality remains the same. Many would benefit from the use of this approach, local governments, universities, medical Institutions and finally the public.

REFERENCES

[1] LADOT. (n.d.). *LA Department of Transportation*. Retrieved 05 2011, from LADOT: http://ladot.lacity.org/tf_preferential_parking_permit.htm

[2] Maccubbin, R. P., & Hoel, L. A. (2000). Evaluating ITS Parking Management Strategies: A Systems Approach. *Virginia Department of Transportation*.

[3] Omnitec Security Dubai UAE. (n.d.). *Security System Solution*. Retrieved 5 2011, from http://www.omnitecsecurity.com/productsview.php?pid=201&catid=2&sp=14

[4] Easy Park Parking Solutions. (n.d.). *Parking Boy Model Parking*. Retrieved 5 2011, from Be/barmatic Parksysteme: http://easypark.biz/easypark/Main/Products-ParkingBoy.htm

[5] Parkmobile International (Australia) Pty Ltd. (2007). *Parkmobile*. Retrieved 5 2011, from Parkmobile Parking Made Simple: http://au.parkmobile.com/

[6] S & K Car Park Management Pty Ltd. (n.d.). *Secure Parking - Home*. Retrieved 5 2011, from SecureParking.com.au: http://www.secureparking.com.au/

[7] San Francisco Municipal Transportation Agency (SFMTA). (n.d.). *SFPark.* Retrieved 6 2011, from SFPark: http://sfpark.org/

[8] Metric Parking. (n.d.). *Metric Parking - AURA.* Retrieved 5 5, 2011, from Metric Parking: http://www.metricparking.com/aura_metricparking.html

[9] APARC. (n.d.). *Vehicle Sensors | APARC Systems*. Retrieved 5 2011, from http://aparcsystems.com/products/vehicle-sensors

[10] The World Bank. (n.d.). *World Development Indicators*. Retrieved 5 2011, from data.worldbank.org: http://data.worldbank.org/data-catalog/world-development-indicators?cid=GPD_WDI

[11] Xinhua. (2011, 9 17). *Number of cars in China hits 100m*. Retrieved 10 2011, from ChinaDaily.com.cn: http://www.chinadaily.com.cn/bizchina/2011-09/17/content_13725715.htm

[12] Nestca. (n.d.). *Car Parking Exchange Platform*. Retrieved 5 2011, from http://www.nestca.com/

[13] Wibur Smith Associates; Michael Baker Jr., Inc. (2009). *FEASIBILITY STUDY OF REAL TIME PARKING INFORMATION AT METRORAIL PARKING FACILITIES (VIRGINIA STATIONS).* Washington Metropolitan Area Transit Authority, Virginia Department of Transportation, Virginia.

[14] World Bank. (2010). *Population, total | Data*. Retrieved 5 2011, from http://data.worldbank.org/indicator/SP.POP.TOTL

[15] World Bank. (2010). *GDP (current US$) | Data*. Retrieved 5 2011, from http://data.worldbank.org/indicator/NY.GDP.MKTP.CD/countries

[16] Wang, M., Huo, H., Johnson, L., & He, D. (2006). *Projection of Chinese Motor Vehicle Growth, Oil Demand, and CO2 Emissions through 2050.* Argonne National Laboratory, Energy Systems Division, Oak Ridge.

INTEGRATED FACILITY LOCATION AND TRANSPORTATION MODELS FOR ROUTING DISASTER RELIEF GOODS

Jae-Dong Hong, Industrial Engineering Technology, South Carolina State University, Orangeburg, South Carolina 29117, (803)536-8861, jhong@scsu.edu
Yuanchang Xie, Civil & Environmental Engineering, University of Massachusetts Lowell, Lowell, Massachusetts 01854, (978)934-3681, yuanchang_xie@uml.edu
KiYoung Jeong, Engineering Management & Systems Engineering, University of Houston at Clear Lake, Houston, TX 77058, (281)283-3862, jeongk@uhcl.edu

ABSTRACT

After emergency events, it is critical through emergency response facilities to distribute emergency supplies for rapid recovery to the affected areas in a timely and efficient manner. The facilities considered in this paper include distribution warehouses, Commodity Distribution Points, and neighborhood locations in need of relief goods. We formulate an integrated model as a mixed integer quadratic programming, construct an Excel spreadsheet model, and solve it by CPLEX for Excel Add-In. Based on this model, we propose two robust models to identify the best facility locations and transportation plans, assuming that some critical infrastructure might be unavailable after a disaster.

Keywords: Emergency Response, Facility Location, Spreadsheet Model, Robust Models

INTRODUCTION

Preliminaries

Several recent weather events inflicted major damage in the populated Eastern Unites States from North Carolina to Vermont. Hurricane Irene left extensive flood and wind damage and rainfall from the remnants of Tropical Storm Lee has eventually forced President Obama to issue a disaster declaration for nineteen (19) Pennsylvania counties. These weather-related emergencies have brought issues of natural disaster planning again. Indeed, after emergencies, it is critical through emergency response facilities to distribute emergency supplies to the affected areas in a timely and efficient manner for rapid recovery. The emergency response facilities include (i) distribution warehouses (DWHs), where emergency relief goods are stored, (ii) intermediate response facilities termed Commodity Distribution Point (CDP) facilities or Disaster Recovery Centers (DRCs), where people can more effectively gain access to relief goods, and (iii) neighborhood locations in need of relief goods. These facilities should be utilized as supply warehouses or distribution centers. The distribution of emergency supplies from the facilities to the affected areas must be done via a transportation network.

Given the significance of transportation costs and the time involved in the distribution of relief goods, the importance of optimally locating warehouses and CDPs is apparent. Traditional facility location models such as set covering models, *p*-center models, *p*-median models, and fixed charge facility location problems (Daskin [1]) implicitly assume that all emergency

response facilities will always operate or be available. However, it is very likely that some emergency response facilities may be damaged or completed destroyed and cannot provide the expected services. When this happens, the demands of the affected areas will have to be satisfied by other facilities that are located farther away than the initially assigned facilities. This obviously will increase the distribution cost and transportation time in supplying relief goods to the affected areas. Compared to the prior-disaster transportation costs minimized by the traditional facility location models, the actual or post-disaster transportation costs can be substantially higher. Thus, it is very important to take into account the post-disaster costs as well as the prior-disaster costs in emergency response facility location modeling. In light of the significant difference in siting between emergency response facilities and other types of facilities and the paucity of the research literature in this area, we propose a new emergency response facility location and transportation model that can better account for the uncertainty caused by the disruptions of critical infrastructure. Assuming that some DWHs might be unavailable after disastrous events, we compare our proposed models based on case studies to demonstrate the developed model's capability to better deal with the risks in emergency response caused by the disruptions of critical infrastructure.

Mathematical Programming Approach

Let *M* be the set of all neighborhoods and potential distribution warehouse (DWH) locations, indexed by *m*. We separate *M* into two sets: *M={N, I}*, where *I* denotes the set of potential distribution warehouse (DWH) locations (indexed by *i =1, 2, ...,w*) and *N* represents the set of neighborhoods (indexed by *n* =1, 2, …, *p*). In this research, we assume CDPs can be located at any neighborhoods and potential DWH locations, while DWH can be built at candidate DWH locations only, since DWH locations must satisfy realistic requirements. Based on these two assumptions, let *J* be the set of potential CDP locations indexed by $j \in \{M\}$, where *j = 1, 2, ...p, p+1, p+2, ...p+i, ...p+w*. Given this problem setting, we formulate the following integer quadratic programming (IQP) model that minimizes the total logistics cost, which is the sum of fixed facility costs and the transportation costs from DWHs to CDPs, between CDPs and neighborhoods (including candidate DWH locations that are not selected), and from DWHs to neighborhoods directly.

$$Minimize\ Z = \left[\sum_{i\in I} a_i W_i + \sum_{i\in I}\sum_{j\in M}\left(\sum_{m\in M} D_m y_{jm}\right) d_{ij} x_{ij}\right] + \left[\sum_{j\in M} b_j B_j + \sum_{j\in M}\sum_{m\in M} D_m d_{jm} y_{jm}\right] + \left[\sum_{i\in I}\sum_{m\in M} D_m d_{im} x_{im}^{d}\right] \quad (1)$$

subject to

$$\sum_{i\in I} W_i \le D_W, \quad (2)$$

$$W_i + B_{p+i} \le 1, \qquad \forall i \in I \quad (3)$$

$$W_i + \sum_{j\in M} y_{j(p+i)} + \sum_{g\in I} x_{g(p+i)}^{d} = 1, \quad \forall i \in I \quad (4)$$

$$\sum_{j\in M} y_{jn} + \sum_{i\in I} x_{in}^{d} = 1, \qquad \forall n \in N \tag{5}$$

$$W_i\, k_i \le \sum_{j\in M} x_{ij} \le W_i K_i, \quad \forall i \in I \tag{6}$$

$$\sum_{i\in I} x_{ij} = B_j, \qquad \forall j \in M \tag{7}$$

$$\sum_{j\in M} B_j \le D_B, \tag{8}$$

$$y_{jm} \le B_j, \quad \forall j \; and \; \forall m \in M \tag{9}$$

$$B_j \cdot L_j \le \sum_{m\in M} y_{jm} \le B_j \cdot U_j, \qquad \forall j \in M \tag{10}$$

$$\sum_{m\in M} D_m y_{jm} \le B_j D_j^{max}, \qquad \forall j \in M \tag{11}$$

$$\sum_{j\in M}\sum_{m\in M} D_m y_{jm}\, x_{ij} + D_i W_i + \sum_{m\in M} D_m x_{im}^{d} \le W_i D_i^{max}, \qquad \forall i \in I \tag{12}$$

$$\sum_{m\in M} x_{im}^{d} \le H_i \cdot W_i, \qquad \forall i \in I \tag{13}$$

where,

$a_i -$ *Fixed cost for contructing and operating DWH* i;
$b_j -$ *Fixed cost for contructing and operating CDP* j;
$B_j -$ 1 *if neighborhood* j *is selected as a CDP*, 0 *otherwise* (*decision variable*);
$d_{ij} -$ *Distance between DWH* i *and CDP* j;
$d_{im} -$ *Distance between DWH* i *and location* m;
$d_{jm} -$ *Distance between CDP* j *and location* m;
$D_B -$ *Maximum number of CDPs can be built* (*set to* 5);
$D_i^{max} -$ *Capacity of DWH* i (500 *for each DWH in this study*);
$D_j^{max} -$ *Capaciy of CDP* j (*set to* 250 *for each CDP in this study*);
$D_m -$ *Demand of location* (*can be either neighborhood or DWH*) m;
$D_W -$ *Maximum number of DWHs can be built* (*set to* 3 *in this study*);
$H_i -$ *Maximum number of locations DWH* i *can directly handle* (*up to* 3);
$k_i -$ *Minimum number of CDPs a DWH must handle* (*set to* 1 *in this study*);
$K_i -$ *Maximum number of CDPs a DWH can handle* (*set to* 5 *in this study*);
$L_j -$ *Minimum number of neighborhoods a CDP needs to cover* (*set to* 2);
$U_j -$ *Maximum number of neighborhoods a CDP can cover* (*set to* 6);
$W_i -$ 1 *if a candidate warehouse* i *is selected*, 0 *otherwise* (*decision variable*);

x_{ij} $-1\ if\ CDP\ j\ is\ covered\ by\ DWH\ i, 0\ otherwise\ (decision\ variable)$;
x_{im}^{d} $-1\ if\ location\ m\ is\ covered\ by\ DWH\ i, 0\ otherwise\ \ (decision\ variable)$;
y_{jm} $-1\ if\ location\ m\ is\ covered\ by\ CDP\ j, 0\ otherwise\ (decision\ variable)$.

Since the main purpose of this paper is to demonstrate how the proposed model works, we further simplify the objective function by excluding the fixed cost terms for CDPs and for DWH. Also, the numbers of CDPs and DWHs to be built are pre-specified. For real-world applications, once the real data are available, such restrictions can be readily relaxed to generate meaningful results. In this paper, we use the following simplified objective function for the simultaneous optimization of DWH and CDP locations.

$$Minimize\ Z = \left[\sum_{i\in I}\sum_{j\in M}\left(\sum_{m\in M} D_m y_{jm}\right) d_{ij} x_{ij}\right] + \left[\sum_{j\in M}\sum_{m\in M} D_m d_{jm} y_{jm}\right] + \left[\sum_{i\in I}\sum_{m\in M} D_m d_{im} x_{im}^{d}\right] \quad (14)$$

Hereafter, this newly introduced model given by Equations (2)-(14) is referred to as **the Regular Facility Location** (RFL) model.

DEVELOPMENT OF ROBUST MODELS

The plan generated by the RFL model is the optimal but may not be optimal after disastrous events. If a DWH becomes unavailable after the disaster, CDPs assigned to this DWH need to be reassigned to other adjacent DWHs with extra capacity. Then the post-disaster logistics cost may become much larger than the pre-disaster optimal cost. To reduce post-disaster logistics cost, one potential solution is to require each CDP to be covered by a backup DWH as well as a main DWH. To do that, we solve the IFL model after changing the right-hand-side of Equation (4) to be 2 from 1 and find the optimal DWH and CDP locations, denoted by $W_i^{*(2)}\ and\ B_j^{*(2)}$. We call this model **the Robust Integer Facility Location** (RIFL) model. Note that the robust model would minimize the post-disaster cost, not the pre-disaster cost. To find the pre-disaster cost for the RIFL model, we solve for the optimal coverage of CDPS and neighbors, $x_{ij}^{*}\ and\ y_{jm}^{*}$, after setting the RHS of Equation (4) back to be 1, with the $W_i^{*(2)}\ and\ B_j^{*(2)}$ fixed.

An alternative way of developing the robust model is to add the capacity constraints of candidate DWHs in a disaster-prone area. For instance, if a DWH has a high probability of being damaged in disastrous events, one can specify that all CDPs assigned to this DWH can only have up to certain percentages of their demand satisfied by it. This strategy would avoid putting all eggs in one basket and improve the robustness of the model. In fact, if a DWH is partially damaged due to disaster, this model would be useful. Now, let x_{ij} be a continuous decision variable between 0 and 1 denoting the fraction of CDP j;s demand satisfied by DWH i. Then, the following capacity constraint is added to the RFL model:

$$0 \leq x_{ij} \leq C_i \cdot B_j, \qquad \forall i \in I \text{ and } \forall j \in M \tag{15}$$

where C_i denotes the maximum fraction of CDP's demand that can be satisfied by DWH i. For candidate DWHs with a high probability of damage or shutdown during disastrous events, C_i would take relatively smaller values, whereas for DWHs in stable and safe areas, C_i would take larger values. By making x_{ij} a continuous decision variable, the robust facility location model becomes a mixed integer quadratic programming (MIQP) problem, which can be linearized by defining a new decision variable as follows:

$$z_{ijm} = x_{ij} \cdot y_{jm}, \tag{16}$$

where z_{ijm} denotes the fraction of neighborhood m's demand satisfied by DWH i via CDP j. Then solving this robust facility location problem is equivalent to solving the problem of the RFL model with (15), after substituting (16) into (14) and (12). We call this model the **Robust Continuous Facility Location** (RCFL) model.

CASE STUDY AND OBSERVATIONS

The regular and robust models can be solved by a variety of optimization software packages, such as LINDO, LINGO, or GAMS. However, coding the developed MILP model using these tools may not be an easy task, since so many decision variables and constraints are involved. Recently, Microsoft Excel spreadsheet-based optimization modeling has become so popular, since spreadsheet models lay out the relevant data, measures of performance, interrelationships, and so forth, on a spreadsheet in an organized way that facilitates fruitful analysis. In this paper, a CPLEX for Microsoft Excel Add-In is used to solve the proposed MILP model. This Excel Add-In provides a friendly user interface that makes coding the MILP model straightforward.

To evaluate the developed MILP model, we conduct a case study using the same cities in South Carolina that Hong et al. [2] used. Twenty (20) cities are selected as neighborhoods and five (5) cities among neighborhoods, with Charleston, Columbia, Florence, Greenville, and Orangeburg, considered as candidate sites for DWHs, as shown in Figure 1. All neighborhoods including five candidate sites for DWHs are candidate locations for CDPs. Table 1 shows the distances (in miles) between any two neighborhoods. Also shown in Table 1 are the demands (in thousands) for all neighborhoods. These demands are hypothetical values proportional to each neighborhood's Year 2000 population and can be readily replaced by true demand data for real-world applications. Based on these input data, an Excel Spreadsheet model is developed.

Using CPLEX for Excel Add-In, we solve the three models, RFL, RIFL, and RCFL for H_i = 0, 1, 2, and 3, for all I, where H_i denotes the maximum number of neighborhoods that DWH i directly covers. To show how robust the RIFL and RCFL models are, two scenarios are considered. The first (normal) scenario assumes that all selected DWHs remain available after disastrous events, whereas the second considers the shutdown/unavailability of a DWH. Hereafter, these scenarios are referred to as normal and shutdown scenarios, respectively. For normal scenario, we evaluate the three models and present the results of facility location and transportation scheme as shown in Table 2 for H_i= 0. From the results under normal scenario, we see that all three models include Columbia and Charleston as selected DWHs. Thus, it would be interesting to see what

would happen if one of DWHs is unavailable and to compare the post-disaster costs of the three models. We select DWH Columbia to be unavailable after disaster, evaluate the three models, and present the results also in Table 2 under the shutdown scenario. We observe that type of robust models affects the optimal CDP locations and transportation scheme. It is interesting to note the changes of covering schemes by DWHs and CDPs under shutdown scenario. For example, in the RFL model under normal scenario, Orangeburg is assigned to CDP Sumter, which is covered by DWH Columbia. Now with DWH Columbia unavailable under shutdown scenario, Orangeburg is now assigned to CDP Beaufort, which is covered by DWH Charleston.

To further investigate the effects of the shutdown of DWHs and to see the performance of the robust models, we consider various shutdown scenarios, present the resulting costs for both scenarios in Table 3, and compare the results for the three models. As expected, the total transportation cost (TTC) for each model increases under the shutdown scenario. As H_i increases, TTC under both scenarios for RFL model always decrease, whereas TTC for two robust models does not always decrease. We observe that, on average, two robust models, RIFL and RCFL, outperform than the non-robust RFL model under the shutdown scenario, though they underperform under the normal scenario.

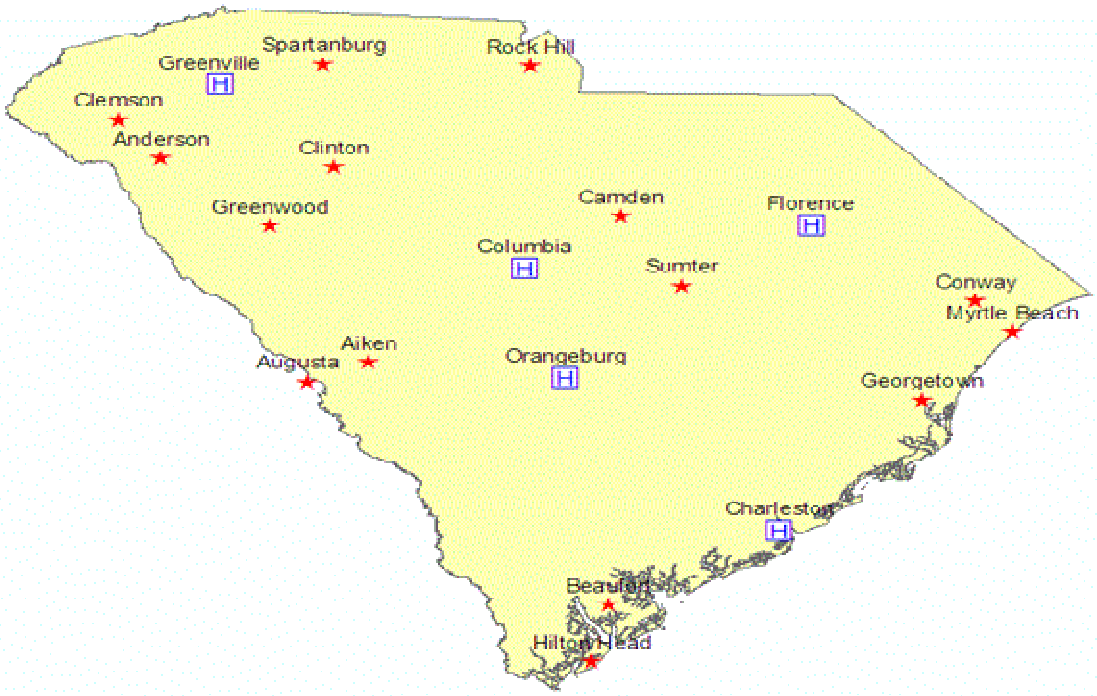


Figure 1. Candidate Warehouses, CDPs, and Neighborhoods

Now, we propose a performance measure index, which is called a robust index (RI), to show how much the results from each model are robust enough to cover the diverse scenarios in terms of cost minimization. Although there are many definitions of robustness, we adopt the one from Dong [3] as "the extent to which the network is able to perform its function despite some damage done to it, such as the removal of some of the nodes and/or link in a network." In the present paper, each model's performance may be evaluated by comparing it with the best performing

model in terms of average TTC and its standard deviation. Hence we propose the following robustness index (RI) for a model g as follows:

$$RI(g) = \alpha \frac{\min_{all\ \lambda} AVG(\lambda)}{AVG(g)} + (1-\alpha)\frac{\min_{all\ \lambda} STD(\lambda)}{STD(g)}, \quad if\ AVG(g) \neq 0; STD(g) \neq 0 \quad (17)$$
$$= \frac{\min_{all\ \lambda} AVG(\lambda)}{AVG(g)}, \quad if\ STD(g) = 0,$$

where AVG(λ) and STD(λ) stand for average and standard deviation of each model λ's cost under given scenarios and α denotes the weight between the average and the standard deviation. Note that as RI for the model becomes closer to 1, the more robust the model would be. And RI can be used to decide the rank of each model in terms of robustness. We calculate RI for the three models for three possible shutdown scenarios and present them in Table 3. We calculate three different RIs- RI for normal scenario and RI for shutdown scenario, and an overall RI for both scenarios with the assumption that all individual scenarios have the same weight. To see the effect of α, the weight between the average and the standard deviation, we calculate overall RI for α = 0, 0.1, 0.2, …, 1 and present them in Table 4. Table 4 supports what we observe from Table 3, that is, the RIFL outperforms a non-robust model, RFL, for α ≠ 1, and the RCFL always surpasses the other two models for all values of α, if the shutdown scenario is included. From Tables 3 and 4, we recommend that the proposed robust models, RIFL and RCFL, be used for optimally locating DWHs under the risk of disruptions. Especially, the RCFL model would work best if the DWH under the scenario has a high probability of being shut down. As discussed previously, transport of relief goods happens mostly after disaster. Therefore, for siting emergency response facilities, it would be more important to minimize the post-disaster cost rather than the pre-disaster cost and to better consider the unavailability of emergency facilities. The example provided here demonstrates that the proposed robust facility location models can well suit the needs of siting emergency response facilities.

SUMMARY, CONCLUSIONS, AND FUTURE RESEARCH

In this paper, we develop an RFL (Regular Facility Location) model and propose two robust transportation and facility location models and compare them with a non-robust RFL. For the RCFL (Robust Continuous Facility Location) model, we introduce a continuous variable, defined in Equation (15), to denote the capacity constraint on a candidate DWH in disaster-prone areas, so that it can only partially satisfy the demand of CDPs, assuming that that each DWH carries limited inventory of emergency relief goods and each CDP also has a finite capacity of storage. We formulate the problem as a mixed integer linear programming model and solve it using CPLEX for Microsoft Excel Add-In. For the RIFL (Robust Integer Facility Location) model, we set the constraint requiring each CDP to be served by multiple DWHs (two DWHs in this paper) on the RFL model, which requires each CDP to be served by one DWH. We propose a performance measure index to show how well the models perform after disaster, RI, defined in (33). Using numerical examples, we show that the two robust models, RIFL and RCFL, yield emergency response facility location plans of slightly higher TTCs (total transportation cost) than the RFL model under normal situations. However, they generate more robust facility location plans in the sense that they can perform better when some of the selected DWHs are

shut down after disaster and these unavailable DWHs can't distribute emergency supplies to the affected areas.

The purpose of establishing emergency response facilities is for distributing relief goods after disaster. Therefore, when evaluating the efficiency and robustness of emergency response facility location plans, more weight should be given to their post-disaster performance. The resulting RIFL and RCFL models are designed in a robust manner such that they can better address scenarios with failures of key transportation infrastructure. Case studies are conducted to demonstrate the developed models' capability to deal with uncertainties in transportation networks. Thus, the developed robust models can help federal and local emergency response officials develop efficient and robust disaster relief plans.

For future research, it would be necessary to develop a robust as well as an integrated transportation and facility location model when both a DWH and a CDP could be unavailable in the shutdown scenario.

REFERENCES

[1] Daskin, M.S. Network and Discrete Location: Models, Algorithms, and Applications. New York, NY: John Wiley, 1995.
[2] Hong, J., Xie,Y. & Jeong, K.Y. *Comparison of Robust Emergency Response Facility Location Models Under the Risk of Disruption.* Proceedings of the 41st Southeast Decision Sciences Institute, 2011, Savannah, GA, February 23-25,800-811.
[3] Dong, M. *Development of supply chain network robustness index*. International Journal of Services Operations and Informatics, 2006, 1(1/2), 54-66.

Table 1(a). Distances (in miles) between Neighborhoods

No.	Neighbor-hoods	Aiken	Anderson	Augusta	Beaufort	Camden	Clemson	Clinton	Conway	George-town	Green-wood	Hilton Head
1	Aiken	0.00	99.69	16.98	121.37	86.19	120.42	69.85	186.02	206.74	55.53	152.40
2	Anderson	99.69	0.00	92.34	246.70	148.32	18.05	50.04	253.07	269.41	39.50	277.66
3	Augusta	16.98	92.34	0.00	127.63	128.68	110.82	81.00	228.30	224.91	62.00	158.59
4	Beaufort	121.37	246.70	127.63	0.00	166.79	271.49	181.15	188.83	137.08	167.60	41.02
5	Camden	86.19	148.32	128.68	166.79	0.00	169.48	87.01	110.14	113.48	102.90	198.23
6	Clemson	120.42	18.05	110.82	271.49	169.48	0.00	63.95	264.79	247.81	56.53	239.42
7	Clinton	69.85	50.04	81.00	181.15	87.01	63.95	0.00	190.71	226.23	26.97	196.77
8	Conway	186.02	253.07	228.30	188.83	110.14	264.79	190.71	0.00	36.62	218.67	193.54
9	Georgetown	206.74	269.41	224.91	137.08	113.48	247.81	226.23	36.62	0.00	247.64	157.04
10	Greenwood	55.53	39.50	62.00	167.60	102.90	56.53	26.97	218.67	247.64	0.00	183.21
11	Hilton Head	152.40	277.66	158.59	41.02	198.23	239.42	196.77	193.54	157.04	183.21	0.00
12	Myrtle Beach	207.12	266.99	225.29	202.69	124.06	262.91	204.74	14.03	34.76	232.70	191.40
13	Rock-Hill	124.47	120.98	142.64	206.76	71.32	120.00	65.57	186.15	232.88	89.97	210.80
14	Spartanburg	142.14	60.36	160.32	225.30	125.80	59.19	35.54	223.24	258.84	59.39	231.61
15	Sumter	112.39	172.26	130.57	125.70	29.34	168.17	104.57	80.81	79.19	116.18	138.17
16	Charleston	162.96	226.73	207.56	70.32	146.74	248.36	170.50	97.41	60.92	191.91	104.98
17	Columbia	56.41	116.50	75.10	134.16	34.69	128.22	61.20	140.20	123.04	72.81	142.64
18	Florence	132.44	192.92	136.00	150.80	50.43	201.61	137.72	53.11	68.54	165.08	170.49
19	Greenville	150.96	31.00	120.94	234.12	134.62	30.10	41.61	231.03	266.62	51.09	234.53
20	Orangeburg	53.75	135.02	76.00	83.91	62.98	161.39	97.82	124.74	105.96	95.52	102.33

Table 1(b). Distances (in miles) between Neighborhoods (continued)

No.	Neighbor-hood (M)	Myrtle Beach	Rock-Hill	Spartan-burg	Sumter	Charleston	Columbia	Florence	Greenville	Orangeburg	Demand (in 1000s)
1	Aiken	207.12	124.47	142.14	112.39	162.96	56.41	132.44	150.96	53.75	29
2	Anderson	266.99	120.98	60.36	172.26	226.73	116.50	192.92	31.00	135.02	26
3	Augusta	225.29	142.64	160.32	130.57	207.56	75.10	136.00	120.94	76.00	196
4	Beaufort	202.69	206.76	225.30	125.70	70.32	134.16	150.80	234.12	83.91	13
5	Camden	124.06	71.32	125.80	29.34	146.74	34.69	50.43	134.62	62.98	8
6	Clemson	262.91	120.00	59.19	168.17	248.36	128.22	201.61	30.10	161.39	12
7	Clinton	204.74	65.57	35.54	104.57	170.50	61.20	137.72	41.61	97.82	9
8	Conway	14.03	186.15	223.24	80.81	97.41	140.20	53.11	231.03	124.74	12
9	Georgetown	34.76	232.88	258.84	79.19	60.92	123.04	68.54	266.62	105.96	9
10	Greenwood	232.70	89.97	59.39	116.18	191.91	72.81	165.08	51.09	95.52	23
11	Hilton Head	191.40	210.80	231.61	138.17	104.98	142.64	170.49	234.53	102.33	48
12	Myrtle Beach	0.00	200.16	237.25	94.56	97.34	146.75	67.14	244.49	138.49	32
13	Rock-Hill	200.16	0.00	61.93	87.32	186.88	67.33	96.09	89.80	108.05	72
14	Spartanburg	237.25	61.93	0.00	130.47	205.42	93.13	170.14	29.09	129.92	37
15	Sumter	94.56	87.32	130.47	0.00	106.14	43.41	39.28	150.20	56.99	41
16	Charleston	97.34	186.88	205.42	106.14	0.00	114.54	109.92	214.24	75.98	121
17	Columbia	146.75	67.33	93.13	43.41	114.54	0.00	79.49	100.91	40.83	130
18	Florence	67.14	96.09	170.14	39.28	109.92	79.49	0.00	177.93	90.34	38
19	Greenville	244.49	89.80	29.09	150.20	214.24	100.91	177.93	0.00	137.71	62
20	Orangeburg	138.49	108.05	129.92	56.99	75.98	40.83	90.34	137.71	0.00	13

Table 2. Results Comparison for Normal/Shutdown Scenarios for Three Models ($H_i = 0$)

Model	**RFL**		**RIFL**		**RCFL**	
Scenario	**Normal**	**Shutdown**	**Normal**	**Shutdown**	**Normal**	**Shutdown**
DWH Selected	1. Charleston 2. Columbia 3. Greenville	1. Charleston 2. Columbia (shutdown) 3. Greenville	1. Charleston 2. Columbia 3. Orangeburg	1. Charleston 2. Columbia (shutdown) 3. Orangeburg	1. Charleston 2. Columbia 3. Greenville	1. Charleston 2. Columbia (shutdown) 3. Greenville
CDPs covered by (DWH #)	1.Beaufort (1) 2.Georgetown (1) 3. Aiken (2) 4. Sumter(2) 5. Spartanburg (3)	1. Beaufort (1) 2. Georgetown (1) 3. Aiken(3) 4. Sumter(1) 5. Spartanburg (3)	1. Beaufort (1) 2. Camden(2) 3. Sumter (2) 4. Clinton (2) 5. Aiken (3)	1. Beaufort (1) 2. Camden(3) 3. Sumter (1) 4. Clinton (3) 5. Aiken (3)	1. Beaufort (1) 2. Georgetown(1) 3. Sumter (2) 4. Greenwood (3) 5. Spartanburg (3)	1. Beaufort (1) 2. Georgetown(1) 3. Sumter (1) 4. Greenwood (3) 5. Spartanburg (3)
Neighbor-hoods Assigned to (CDP)	•(Beaufort), Hilton-Head •(Georgetown), Conway Myrtle-Beach •(Aiken), Augusta, •(Sumter), Camden , Florence, Orangeburg •(Spartanburg), Anderson, Clemson, Clinton, Rock-Hill, Greenwood	•(Beaufort), Hilton-Head, Orangeburg •(Georgetown), Conway, Myrtle-Beach, Florence •(Aiken),August •(Sumter), Camden Columbia •(Spartanburg), Clinton Anderson, Rock-Hill, Clemson, Greenwood	•(Beaufort), Hilton-Head •(Camden), Rock-Hill •(Sumter), Conway, Florence, Georgetown, Myrtle-Beach •(Clinton),Anderson, Clemson, Greenwood, Spartanburg, Greenville, •(Aiken), Augusta	•(Beaufort), Hilton-Head •(Camden), Conway, Rock-Hill •(Sumter), Columbia, Florence, Georgetown, Myrtle-Beach •(Clinton), Anderson, Clemson, Spartanburg, Greenville •(Aiken), Augusta, Greenwood	•(Beaufort), Hilton-Head •(Georgetown), Conway, Myrtle-Beach •(Sumter), Camden, Florence, Orangeburg •(Greenwood), Aiken, Augusta •(Spartanburg), Anderson, Clemson, Clinton, Rock-Hill	•(Beaufort), Hilton-Head, Orangeburg •(Georgetown), Conway, Myrtle-Beach, Florence •(Sumter),Camden, Columbia •(Greenwood), Aiken, Augusta •(Spartanburg), Anderson, Clemson, Clinton, Rock-Hill
(CDB,CBN, CDN) TTC	($29,759, $17,741, $0) $47,500 (A)	($68,919, $24,847, $0) $93,766 (B)	($68,919, $24,847, $0) $58,448 (A)	($64,230, $29,869, $0) $94,099 (B)	($29,068, $26,809, $0) $55,877 (A)	($46,955, $33,914, $0) $80,869 (B)
Increase (B)-(A)	$46,266		$35,651		$24,992	

CDB: Cost from DWHs to CDPs, 1st Term in Eq. (14). **CBN**: Cost from CDPs to Neighbors, 2nd Term in Eq. (14).
CDN: Cost from DWHs to Neighborhoods, 3rd Term in Eq. (14). **TTC**= CDB + CBN + CDN

Table 3. Comparison Between Integrated and Two Robust Models

Shutdown Scenario	Model ($H_i = 0$)					
	RFL		**RIFL**		**RCFL**	
	Normal	**Shutdown**	**Normal**	**Shutdown**	**Normal**	**Shutdown**
DWH 1	$47,500	$93,766	$58,933	$94,099	$55,877	$80,869
DWH 2	$47,500	$93,630	$58,933	$78,424	$47,953	$81,063
DWH 3*	$47,500	$77,170	$58,933	$81,304	$47,500	$76,259
AVG	**$47,500**	**$88,188**	**$58,933**	**$84,609**	**$50,443.**	**$79,397**
STD	**0**	**$9,542**	**0**	**$8,343**	**$4,711**	**$2,719**
RI	**1**	**0.705**	**0.806**	**0.761**	**0.471**	**1**
Shutdown Scenario	($H_i = 1$)					
	Normal	**Shutdown**	**Normal**	**Shutdown**	**Normal**	**Shutdown**
DWH 1	$44,572	$85,096	$46,385	$77,344	$46,594	$74,798
DWH 2	$44,572	$81,795	$46,385	$73,215	$46,311	$71,154
DWH 3	$44,572	$75,021	$46,385	$72,092	$48,133	$69,344
AVG	**$44,572**	**$80,637**	**$46,385**	**$74,217**	**$47,012**	**$71,765**
STD	**0**	**$5,136**	**0**	**$2,765**	**$980**	**$2,778**
RI	**1**	**0.693**	**0.961**	**0.943**	**0.474**	**1**
Shutdown Scenario	($H_i = 2$)					
	Normal	**Shutdown**	**Normal**	**Shutdown**	**Normal**	**Shutdown**
DWH 1	$43,058	$79,163	$44,917	$78,593	$44,822	$74,657
DWH 2	$43,058	$73,386	$44,917	$75,530	$43,676	$70,151
DWH 3	$43,058	$73,672	$44,917	$69,203	$44,324	$68,468
AVG	**$43,058**	**$75,407**	**$44,917**	**$74,442**	**$44,274**	**$71,092**
STD	**0**	**$3,256**	**0**	**$4,788**	**$574**	**$3,200**
RI	**1**	**0.922**	**0.958**	**0.784**	**0.486**	**1**
Shutdown Scenario	($H_i = 3$)					
	Normal	**Shutdown**	**Normal**	**Shutdown**	**Normal**	**Shutdown**
DWH 1	$42,671	$77,974	$44,557	$76,567	$43,976	$73,708
DWH 2	$42,671	$73,082	$44,557	$74,095	$45,322	$69,219
DWH 3	$42,671	$72,551	$44,557	$69,211	$43,418	$68,466
AVG	**$42,671**	**$74,535**	**$44,557**	**$73,291**	**$44,238**	**$70,464**
STD	**0**	**$2,989**	**0**	**$3,743**	**$979**	**$2,834**
RI	**1**	**0.911**	**0.957**	**0.831**	**0.482**	**1**
Overall AVG	**$62,071**		**$62,669**		**$59,836**	
Overall STD	**$18,793**		**$15,595**		**$14,178**	
Overall RI	**0.859**		**0.932**		**1**	

(1)AVG and STD stand for average and standard deviation, respectively, (2) Alpha (α) is set to 0.5 for RI. (3)DWH1: Columbia for all models. DWH2: Charleston for all models. DWH3*: Orangeburg for RIFL, Greenville for other models: DWH3: Greenville for all models.

Table 4. Comparison of Overall Robust Index Between Integrated and Two Robust Models

Alpha (α)	RFL	RIFL	RCFL
1.0	0.964	0.955	1.000
0.9	0.943	0.950	1.000
0.8	0.922	0.946	1.000
0.7	0.901	0.941	1.000
0.6	0.880	0.937	1.000
0.5	0.859	0.932	1.000
0.4	0.838	0.927	1.000
0.3	0.817	0.923	1.000
0.2	0.796	0.918	1.000
0.1	0.775	0.914	1.000
0.0	0.754	0.909	1.000

THE EFFECTS OF CODES OF ETHICS ON THE SUPPLY CHAIN: A COMPARISON OF LEs AND SMEs

Barbara Withers, School of Business Administration, University of San Diego, San Diego, CA 92110, (619) 260-2380, bwithers@sandiego.edu
Maling Ebrahimpour, University of South Florida, St. Petersburg, St. Petersburg, FL 33701, (727) 873-4786, bizdean@stpete.usf.edu

Abstract

A study of the perceptions of supply chain management professionals toward CSR was conducted to determine whether exposure to codes of conduct and ethics training influenced their behaviors in the workplace. Survey results revealed that most companies have corporate codes of conduct in place. Having a corporate code of conduct did not have a significant impact on employees' commitment to CSR, but mandatory training on the codes did significantly impact the level of commitment towards CSR. The level of perceived commitment to CSR impacted how often the respondents witnessed various unethical behaviors within their respective companies. Higher levels of inappropriate behaviors were observed among vendors' employees when one's codes of conduct were shared with vendor employees. True commitment from top management makes a significant difference in how employees perceive and respond to codes of conduct and social responsibility.

Keywords: Codes of Ethics, Supply Chain

INTRODUCTION

The issue of unethical and illegal activities in the business environment is one that no business can afford to take lightly. The results of corporate misconduct (intentional and unintentional) have already had dramatic negative impacts on society and on corporate performance. In the late 1990s, Estes [20] estimated that the social costs associated with unethical and socially irresponsible corporate decisions were approximately 2.5 trillion dollars annually. Now, more than a decade later, the costs are inevitably higher.

The need for strong ethical standards and actions is greater now than ever before. Today's global business environment is even more heavily impacted by corruption, high taxes, and intense competition making firms vulnerable to unethical or illegal actions and practices (for example, bribery, inappropriate labor practices, etc.) [62]. In addition, in difficult economic environments, the pressures to survive (individually as well as organizationally) often result in ethical concerns being ignored [85]. To address these concerns, increasing numbers of firms are adopting formal measures to regulate themselves by creating corporate codes of ethics [12] [22] [23].

Codes of ethics are referred to by a variety of names in the literature, including codes of conduct, codes of practice, corporate credos, mission statements, and values statements [65]. Schwartz also synthesized the many definitions of codes used in the literature and defined a code of ethics as "a written, distinct and formal document which consists of moral standards used to guide employee or corporate behavior" [65, p. 248]. For the purposes of this study, Schwartz's definition of code of ethics will be used. Additionally, the terms "code of ethics" and "code of conduct" may be used interchangeably.

Codes of conduct have evolved according to business and societal needs. Toward the beginning of the 20th century, businesses functioned under the assumption that our societal norms and personal values were sufficient to drive appropriate behaviors in organizations. But as organizations (particularly business organizations) became more complex by mid-century, it became evident that people needed more specific guidance on how to deal with the types of situations they were encountering in the work place. As a result, businesses found it necessary to create formal written corporate codes of ethics [7] [44].

In the latter half of the century, with the advent of the computer, the internet, international travel, multi-national corporations, access to global resources, and a host of other events, organizational environments had become so complex that the need for standards to guide behaviors and actions in an international context was apparent. While typical corporate codes addressed the basic ethical issues businesses in the U.S. were likely to encounter (e.g., integrity, confidentiality, and sexual harassment), the international business environment exposed businesses to issues with which firms had little or no experience, for example, human rights, child labor, bribery, quality of life, and sweatshops. These needs called for major revisions or expansions to existing codes of conduct to assure that they could be effective in the international business context [5].

Each decade's ethical scandals seemed to draw attention to another sphere of unethical behaviors. Events like Watergate in the 1970s, corporate raiding in the 1980s, abusive labor practices that dominated the textile and sports wear industries in the 1990s, and financial and accounting debacles around the turn of the century spurred the creation of different types of codes to address different organizational ethical issues. Toward the end of the 20th century, standards by governmental and non-governmental organizations were being created to address basic rights and moral behaviors that would span national and international boundaries and guide the actions and behaviors of people around the world (for example, the International Labour Organization, the World Health Organization, the International Standards Organization (ISO)). Around the same time businesses realized that one set of codes did not necessarily meet the needs of an entire organization, so sub-organization codes began to be developed to address the needs of parent companies and subsidiaries more effectively.

Media attention in the past two decades also revealed that certain industries and professions were susceptible to inappropriate ethical behaviors more frequently than others (for example, the textile industry, the accounting profession, and the purchasing process). These conditions led to the creation not only of industry-wide codes of

conduct (for example, purchasing standards), but also of laws and statues designed to prevent such violations from occurring again (e.g. Sarbanes-Oxley Act). Similar awareness was developing within businesses as they began to recognize that the three business functions (finance, marketing, and operations) faced different situations and contexts and that codes of conduct were needed for these functional areas. The shift toward functional codes is no longer in its infancy, but codes have evolved more quickly for the finance and marketing functions than for the operations function.

The last stage of code evolution, to date, appears to be the growing awareness that codes within organizations (e.g., corporate codes or marketing codes) or as umbrellas over organizations (e.g., industry codes or international standards) cannot lead to the type of sustainable behaviors by all of the relevant to achieve the goals of businesses and society at large. These types of codes, which are only beginning to be created, recognize the inter-related needs of business organizations along the supply chain as well as the needs of the communities and societies affected by the operations of these business organizations. Hence, there is the recognition that codes of conduct for ethical and moral behaviors affect far more than the single organization that may have crafted the code and that these codes must work in concert with each other [57] to produce effective results in terms of appropriate behaviors and decisions that will benefit the triple bottom line (profits, people, and planet).

The evolution of the codes themselves has been mirrored to some degree by the patterns in scholarly research. Earlier business ethics research concentrated on corporate level codes and the corporation's overall responsibility for good citizenship. In the past two decades, research directed to the social obligations and responsibilities of firms increased. Research on sub-organizational codes has finally begun.

Several elements exist that create the basis for the research reported in this study: the increased awareness of the important role that the supply chain plays in meeting corporate social responsibility (CSR) [37] [62] [34] [90], the paucity of ethics research that considers small and medium sized firms [86] [41], and the fact that current research has not established the effectiveness of codes of conduct [46] [79] [65].

The role of the supply chain in determining firms' abilities to achieve their CSR goals has increased dramatically in the past decade [75] [3]. The supply chain function is the obvious one for influencing CSR and ethics initiatives simply because the supply chain has the ability to influence virtually all of the activities that would typically affect socially responsible behaviors.

Very few studies have addressed small to medium-size firms (SMEs) or compared the experiences of SMEs with large firms and most of this work relates to European firms [27] [53] [60]. This is a severe shortcoming in the research for several reasons: 1) the overwhelming majority of firms in the US and Europe are small to medium sized but this majority has been effectively ignore in the research, and 2) the characteristics, structures, and operations of small to medium sized firms are sufficiently different from those of large firms that it is most likely inappropriate to extrapolate from most research and apply

conclusions to small and medium size firms' behaviors. The limited research that specifically addresses small to medium size firms suggests that their behaviors, philosophies, and perceptions are considerably different from those of large firms. This observation needs to be further tested.

Researchers have struggled to determine whether codes of conduct have had the desired impact on corporate decision-making and employee behaviors. Considerable research has already been undertaken to address code effectiveness in large enterprises [24] [48] [2] [6] [9] [79] [59] [19] [35]. Unfortunately the results from these studies have been mixed suggesting that this issue has not been resolved.

The purpose of this study is to compare the effectiveness of codes of conduct for large enterprises (LEs) with the effectiveness of codes in small to medium sized enterprises (SMEs). Code effectiveness will be assessed for the company's own employees as well as for the employees of the company's upward and downward supply chain partners.

METHODOLOGY

Questionnaire Development

Our questionnaire was developed in a six-step process. The first step was a comprehensive review of the literature on ethics and the supply chain. Within this literature review, particular effort was made to identify articles that addressed ethics in the context of CSR. Based on the literature review, the second step was to draft the survey instrument. The third step of the questionnaire development process was an initial test for clarity by graduate students at a major US university. Using information from this test, the questionnaire was modified to improve understandability. The next step was a test for clarity, relevance, and technical accuracy and was administered to members of the Supply Chain Management Institute (SCMI) affiliated with a west coast university. The final step was a refinement of the survey instrument based on the results of the survey responses from these SCMI members.

The survey instrument contained a series of questions to capture demographic information about the respondents and their respective firms. The other questions in the survey were designed to capture information about firms' actions relevant to ethics and social responsibility. Most of these questions utilized a five-point Likert scale with five representing the strongest or most positive response and one representing the weakest or most negative response. Another set of questions was designed to assess observed behaviors to which respondents answered yes (coded as 1) or no (coded as 2). For the purposes of this study, mean responses for both types of questions were analyzed.

Sample Selection and Characteristics

The population for the survey was comprised of supply chain professionals /members of the Institute for Supply Management ISM). A random sample of 5000 members, representing twelve major industries, was provided by ISM from its membership list.

The industries selected were Agriculture; Mining; Utilities; Construction; Manufacturing; Wholesale Trade; Retail Trade; Transportation and Warehousing; Information Services; Finance and Insurance; Professional, Scientific, and Technical Services; Administrative and Support and Waste Management and Remediation Services; Educational Services; Health Care and Social Assistance; Accommodation and Food Services; Other Services (not Public Administration); and Public Administration.

Of the 5000 names and addresses provided by ISM, 63 names were eliminated because of incomplete mailing addresses. Hard copy surveys were mailed to the remaining 4,937 addressees obtained from the sample draw. Nine surveys were returned to sender, reducing the population size to 4,928. Of these, a total of 421 completed and usable survey forms were returned, for a response rate of 8.5%.

Responding firms represent both multinational and US-based companies: 61.1% of the respondents indicated their company is a multinational company and 35.1% indicated their companies are not multinational. A small portion (3.8%) indicated they do not know whether their companies are multinational. The vast majority of the respondents' companies are headquartered in North America (89.7%); 8.6% of respondents' companies are headquartered in Western Europe. The remaining companies have headquarters scattered in Central and South America, Eastern Europe, or Asia. The vast majority of respondents' firms (90.2%) identified North America as their primary sales region; 84.7% indicated that their firms' primary sources of purchases were also from North America.

Survey respondents work in all of the industries represented in the original sample but the majority are employed in two industries: Manufacturing (45.3%) and Utilities (14.2%). Firm size was evaluated in terms of annual sales as well as number of employees. In terms of annual sales, firms with sales less than or equal to $500 million were classified as small to medium (SMEs); 56.3% of the respondents' firms fell into this category. Firms with sales greater than $500 million were classified as large; 35.6% of respondent's firms fell into this category. Just over 8% of respondents did not answer this question. Firm size, as represented by number of employees, was divided into two groups: small to medium sized firms (SMEs) were defined as having 1000 employees or less (27.2% of respondents were in this category), and large firms (LEs) were defined as having more than 1000 employees (72.2% were in this category). Only a half percent of respondents were unable to respond to the question on the number of employees in their organizations.

An analysis of the difference in firm sizes indicated that responses from employees at LEs (as measured by number of employees) were significantly different from the responses from employees at SMEs. As a result, the sample was divided into two sub-sets: LEs and SMEs. Statistical analyses were conducted on each set and are reported below.

FINDINGS

Written codes of ethics are the primary mechanism that firms can use to influence the ethical behaviors of employees [79]. Schwartz [65] maintained that the effectiveness of codes of ethics depends upon the degree to which the code is communicated to employees. Furthermore, studies have also revealed that top management commitment plays a critical role in determining whether codes of ethics will influence employee behaviors [79]. Working from these constructs, the first set of questions in our survey was designed to capture the ethical environment of responding firms.

To determine whether companies recognized the need to have a means for communicating their corporate values to employees and members of their supply chain, respondents were asked to indicate whether their companies have a written policy addressing business ethics and social responsibility. For LEs, 96.7% of the respondents indicated what their companies have written policies in one of the following forms: a code of ethics, a policy statement on ethics, a code of conduct, or a set of guidelines addressing ethical behavior. Only 77.6% of SMEs have written policies on business ethics. ANOVA revealed that the existence of a written code of ethics was significantly influenced by firm size (p=.000).

Schwartz [65] indicated, however, that having a written policy on ethics and social responsibility is a necessary component for influencing the behaviors of employees but is not sufficient if the code is not effectively communicated to employees. Furthermore, if the firm's ethics policy is not perceived as being valued by management, then employees will tend not to embrace the code. Tucker [78] suggested that management can demonstrate commitment to its codes of ethics by investing in their success (i.e., devoting human and financial resources to train employees on the codes).

We used several questions to assess the prevalence of the dissemination of companies' codes of ethics to employees. First, respondents were asked to identify the proportion of employees to whom the company's code of ethics had been communicated. Over 90% of respondents from LEs indicated that their company's code of ethics had been communicated to at least 76% of its employees. In contrast, only 66.4% of respondents from SMEs indicated that their company's code of conduct had been communicated to at least 76% of its employees. This difference is significant (p=.000).

Secondly, when asked whether training was provided on the company's code of ethics, 81.4% of respondents from LEs indicated that code training was mandatory. In contrast, only 39.7% of respondents from SMEs indicated that mandatory training occurred. This difference was significant (p=.000).

Previous studies have shown that the existence of codes of ethics contributes to the corporate culture by creating an ethical environment for employees that is conducive to ethical decision-making [79] [33] [82] [65]. According to Adams [1] firms embrace codes of ethics for a number of reasons but an important one is to institutionalize corporate values into the culture and create an environment that will help socialize

employees into behaviors consistent with this culture. Tucker [78] argued that the existence of codes will not eliminate bad behavior on the parts of employees unless the firm can show that top management is committed to supporting these codes (financially and with human resources). Working from these frameworks, we utilized a set of questions to determine whether the existence of a code of ethics (an attempt to institutionalize corporate values) or training on the code (a significant commitment of human and financial resources) have impacted the perception that employees and managers care committed to social responsibility.

As indicated above the vast majority (96.7%) of large firms have written codes while a smaller majority (77.6%) of small firms have written codes. Perceptions of the level of commitment to social responsibility were obtained using a 5-point Likert scale, with 1 representing "not committed" and 5 representing "completely committed." Separate questions were used to assess perceptions about employee commitment and top management commitment to social responsibility. ANOVA was used to compare the responses to these questions by large firms and small firms.

For LEs, having a written code of ethics has a very significant (p=.000) impact on the perception that top management is committed to social responsibility. A written code also correlated positively to the perception that employees are committed to social responsibility. This difference was significant at the .05 level. Interestingly, for SMEs, the existence of written codes had no impact on perception of either employee or top management commitment to social responsibility. SMEs with written codes were only slightly more likely to perceive employees or top management as committed to social responsibility than were firms without written codes.

Since Schwartz [65] maintained that employee perceptions and code effectiveness would be impacted by the degree of training on the code, we felt it would be more insightful to see if mandatory training on the company's code strengthened the perception of commitment to CSR. The relationship between mandatory training on corporate codes of ethics and perceived commitment to social responsibility was evaluated for each firm size. ANOVA revealed that for LEs, mandatory training has a significant impact (p=.000) on perceptions about both employee and top management commitment to social responsibility. In the case of SMEs, mandatory training did lead to significantly higher perceptions that top management is commitment to social responsibility (p=.01) but had no impact on perceptions about employee commitment.

Next we were interested in whether the existence of a formal written code, the practice of mandatory training, and the perception of employee and top management commitment to social responsibility have affected behavior in the workplace. To assess these impacts, respondents were asked how often they had observed conduct at their company that either violated the law or their company's code of conduct. A 5 point Likert scale with the following options was used for responses: 1 = never observing violating behaviors; 2 = rarely observing behavior violations, 3 = occasionally observing behavior violations but not on a regular basis, and 4 – occasionally observing behavior violations but on a regular basis, and 5 = frequent observations of behavior violations. Respondents were also asked

to indicate the type of illegal or inappropriate behaviors they had witnessed at their companies.

Almost all (96.7%) of LEs have a *written code* of ethics. ANOVA shows that LEs that have a written code of ethics experienced significantly lower levels of behavioral misconduct than did firms without a written code (p=.05). *Mandatory training* on ethics also had a positive impact on behavior. Over 81% of LEs have mandatory training on their codes. ANOVA results comparing the prevalence of mandatory training with the frequency of occurrence of misconduct by employees indicates that mandatory training has a significant beneficial effect on behavior (P=001). Both *top management commitment and employee commitment* to social responsibility impacted how often misconduct was observed in LEs. The stronger the commitment by employees and top management, the lower the frequency of misconduct. Each of these differences was significant (p=.000).

For SMEs, results were not comparable. Recall that a smaller proportion (77.6%) of small firms have a formal written code of conduct or ethics policy. SMEs that do not have a *written code* of ethics experienced only slightly higher levels of misbehavior than was seen in firms with written codes. The difference was not significant. With respect to the impact of training, MEs garnered no beneficial results from *mandatory training* on their codes of ethics. There was no significant reduction in misconduct for SMEs with mandatory training compared with those without mandatory training. And finally, with respect to the impact that *employee and top management commitment* had on the occurrence of misconduct, SMEs also experienced reductions in the frequency of misconduct as a result of employee and top management commitment to CSR. Each of these impacts was significant (p=.05).

Concern about the ethicality of actions by members of a company's supply chain have grown in the years since the sweatshop scandals in the 1990s. Increasingly companies have begun sharing their codes of conduct with others in their supply chains. To assess the prevalence of this practice by the firms represented in this survey, respondents were asked whether a code of conduct has been provided to their suppliers and/or their B2B customers as well as whether these members of the supply chain were expected to adhere to the code. The comparison of LE and SME practices showed that 68% percent of LEs provide vendors with a copy of their respective codes of conduct while only 44.8% of SMEs share their codes with vendors. Interestingly, 83.3% of respondents from LEs expect their suppliers to respect their company's code, even if this code had not been explicitly communicated to the vendor. In contrast, 56.9% of SMEs expect their vendors to respect their codes. Both of these differences are significant (p=.000).

Less sharing of codes exists in the other direction of the supply chain. Among LEs, 44.1% share their code of ethics with B2B customers. In contrast, 29.3% of SMEs share their codes with B2B customers. However, both LEs and SMEs expect higher proportions of their B2B customers to respect corporate codes of ethics during business transactions (63.7% of LEs vs. 41.4% of SMEs). Both differences are significant (p=.000).

The last set of questions in this study attempted to assess whether the company's efforts to create an environment conducive to ethical behavior had an impact on the occurrence of misconduct within the supply chain. Based on the literature, eleven types of misconduct associated with social responsibility and ethical conduct were identified and used in the questionnaire (see Table 1). These behaviors were evaluated to determine whether they had been impacted by the awareness of codes of ethics.

Table 1 – Types of Illegal Conduct or Code Violations

By Company or Vendor Employees:
Abusing drugs or alcohol
Engaging in sexual harassment
Giving or accepting bribes, kickbacks, or inappropriate gifts
Falsifying records and reports
Lying to employees, customers, vendors, or the public
Withholding needed information from customers, vendors, or the public
Misreporting actual time or hours worked
Stealing, theft, or related fraud
Breaking environment and safety laws or regulations
Abusing or intimidating other company employees
Discriminating on the basis of race, color, gender, age, or similar categories
By Customer Employees:
Abusing drugs or alcohol
Engaging in sexual harassment
Giving or accepting bribes, kickbacks, or inappropriate gifts
Falsifying records and reports

To assess how these behaviors may have been impacted, we first evaluated whether sharing one's code of conduct with members of one's supply chain affected behaviors by supply chain employees. ANOVA was used to assess the impacts. For large firms, sharing their codes of conduct with their vendors resulted in significant reductions in only two types of misconduct among vendor employees: giving or accepting bribes, kickbacks, or inappropriate gifts, and stealing, theft, or related fraud ($p=.05$). When large firms shared their codes with B2B customers, there was no significant impact on the behaviors of B2B employees.

For small firms, sharing codes of conduct with vendors had no impact on the behaviors of vendor employees. However, sharing codes with B2B customers significantly reduced the occurrences of discrimination on the basis of race color, gender, age, or similar categories ($p=05$).

Table 2 – Impact of Mandatory Training on Illegal Conduct or Code Violations

Illegal Conduct or Code Violation	Sign. Lg.	Sign. Sm.
By Company Employees:		
Abusing drugs or alcohol		
Engaging in sexual harassment	***	
Giving or accepting bribes, kickbacks, or inappropriate gifts	**	
Falsifying records and reports	****	
Lying to employees, customers, vendors, or the public	**	
Withholding needed information from customers, vendors, or the public		
Misreporting actual time or hours worked	**	
Stealing, theft, or related fraud	***	
Breaking environmental and safety laws or regulations	****	*
Abusing or intimidating other company employees	*	
Discriminating on the basis of race, color, gender, age, or similar categories		
By Vendor Employees:		
Abusing drugs or alcohol		
Engaging in sexual harassment		
Giving or accepting bribes, kickbacks, or inappropriate gifts		
Falsifying records and reports		
Lying to employees, customers, vendors, or the public		
Withholding needed information from customers, vendors, or the public		
Misreporting actual time or hours worked		
Stealing, theft, or related fraud		
Breaking environmental and safety laws or regulations		
Abusing or intimidating other company employees		
Discriminating on the basis of race, color, gender, age, or similar categories		
By Customer Employees:		
Abusing drugs or alcohol		
Engaging in sexual harassment		
Giving or accepting bribes, kickbacks, or inappropriate gifts		****
Falsifying records and reports		

* p=.05
** p=.01
***p=.001
****p=.000

Next, we sought to determine whether mandatory training or employee commitment or top management commitment to social responsibility (a company's internal ethical

environment) have had an impact on the eleven types of misconduct by the company's own employees, vendor employees, or B2B customer employees.

The frequency of mandatory training within respondents' firms was compared to the types of inappropriate behaviors observed among company employees and vendor employees. All eleven types of misconduct were used for these comparisons. For B2B employees, only four of these types of behaviors (abusing drugs or alcohol; engaging in sexual harassment; giving or accepting bribes, kickbacks, or inappropriate gifts; and falsifying records or reports) were considered. The results for both large and small firms are shown in Table 2. Within large firms, having mandatory training on codes of conduct has resulted in significantly lower incidences of illegal conduct or code violations among a company's own employees for eight of the eleven types of misconduct. However, among small firms, the use of mandatory training on corporate codes had little or no impact on employee behaviors. The only behavior significantly affected by mandatory training was "breaking environmental and safety laws or regulations."

The next issue related to determining whether training one's own employees on the company's code of conduct "rubs off" on others outside of the firm, i.e., affects the behaviors of employees in the supply chain. Evidence in the literature indicates that vendor employees modify their behaviors in the presence of customer representatives [19] [90] [91]. The results of these analyses revealed that mandatory training of one's own employees had no impact on the behaviors of vendor employees for either large or small firms. The same was the case for the impact of mandatory employee code training on the observation of misconduct by B2B customer employees for both large and small firms, with one exception. For small firms, mandatory training of company employees had a significant impact ($p=.000$) in reducing the occurrence of "giving or accepting bribes, kickbacks, or inappropriate gifts."

Based on the literature, top management commitment has been shown consistently to be a critical element in the successful implementation of most corporate initiatives, for example, JIT, TQM, ISO 9000 [64] [15] [18] [87] [11] [43]. For this reason, it is assumed that top management commitment to social responsibility would set the tone for establishing an environment perceived to encourage and support ethical behaviors. To see if this is the case, we examined the perception of top management commitment to social responsibility with responses about the occurrences of inappropriate behaviors by employees within the firms as well as by vendor and B2B employees. Results of these analyses are provided in Table 3. Among respondents from large firms, we found that higher top management

commitment to social responsibility resulted in significantly lower incidences of misconduct for seven of the eleven types of illegal conducts or code violations. However, this breadth of impacts did not occur in small firms. Higher levels of top management commitment in small firms did not contribute to reducing employee misconduct except for the "breaking of environmental and safety laws or regulations." For this misconduct, high levels of top management commitment led to significantly lower ($p=.000$) incidences.

Since top management is so important for setting the tone within organizations, we felt it would be interesting to see if the level of top management commitment to social responsibility would influence the occurrence of unethical behaviors in the supply chain. In large firms, top management commitment had a significant beneficial impact on only four types of misconduct (see Table 3). For small firms, top management commitment to social responsibility had no impact of any type of misconduct. Top management commitment had no impact on the behaviors of B2B customers' employees, for either firm size.

The level of employee commitment to social responsibility should manifest itself into appropriate behaviors within the company. Furthermore, it seems reasonable to expect that employees with high levels of commitment to social responsibility would interact with members of the supply chain in ways that would encourage appropriate behaviors on their parts. To see if this is the case, the level of employee commitment to CSR was compared with the incidences of misconduct within the company a well as by employees in the supply chain. Results are shown in Table 4.

Once again, we see evidence that the perception of a more positive ethical environment within large firms seems to manifest itself in improved behaviors by employees. Higher levels of employee commitment to social responsibility significantly reduced the occurrences of most ethically questionable behaviors (See Table 4). Only two types of misconduct (abusing drugs or alcohol and stealing, theft, or related fraud) did not benefit from higher levels of employee commitment to CSR. For small firms, the level of employee commitment to social responsibility had virtually no impact on curbing misbehavior by employees; high levels of employee commitment to CSR significantly reduced only the incidence of breaking environmental and safety laws or regulations ($p=.001$).

In terms of whether employee commitment to CSR has any influence on the behaviors of employees in the supply chain, there appears to have been some "carry over" effect. For large firms, with respect to vendor employees' behaviors, four types of vendor misconduct were significantly reduced as a result of higher employee commitment to CSR: engaging in sexual harassment; giving or accepting bribes, kickbacks, or gifts; falsifying records and reports; stealing, theft or related fraud (see Table 4). For small firms, higher employee commitment reduced only the occurrences of discriminating on the basis of race, color, gender, age, or similar categories to a significant degree ($p=.05$).

Employee commitment to CSR had some impact on B2B customers' employees' behaviors. For large firms, stronger employee commitment to social responsibility was significantly correlated to reductions in engaging in sexual harassment and falsifying records and reports (see Table 4). For small firms, the only beneficial impact of higher employee commitment to CSR was a reduction in the occurrences of falsifying records and reports ($p=.05$).

Table 3 – Impact of Top Management Commitment on Illegal Conduct or Code Violations

Illegal Conduct or Code Violation	Sign. Lg.	Sign. Sm.
By Company Employees:		
Abusing drugs or alcohol	*	
Engaging in sexual harassment		
Giving or accepting bribes, kickbacks, or inappropriate gifts	****	
Falsifying records and reports	*	
Lying to employees, customers, vendors, or the public	****	
Withholding needed information from customers, vendors, or the public	****	
Misreporting actual time or hours worked		
Stealing, theft, or related fraud		
Breaking environmental and safety laws or regulations	**	****
Abusing or intimidating other company employees	****	
Discriminating on the basis of race, color, gender, age, or similar categories		
By Vendor Employees:		
Abusing drugs or alcohol		
Engaging in sexual harassment		
Giving or accepting bribes, kickbacks, or inappropriate gifts	*	
Falsifying records and reports		
Lying to employees, customers, vendors, or the public	**	
Withholding needed information from customers, vendors, or the public		
Misreporting actual time or hours worked		
Stealing, theft, or related fraud	*	
Breaking environmental and safety laws or regulations		
Abusing or intimidating other company employees		
Discriminating on the basis of race, color, gender, age, or similar categories	**	
By Customer Employees:		
Abusing drugs or alcohol		
Engaging in sexual harassment		
Giving or accepting bribes, kickbacks, or inappropriate gifts		
Falsifying records and reports		

*p=.05
**p=.01
***p=.001
****p=.000

Table 4 – Impact of Employee Commitment on Illegal Conduct or Code Violations

Illegal Conduct or Code Violation	Sign. Lg.	Sign. Sm.
By Company Employees:		
Abusing drugs or alcohol		
Engaging in sexual harassment	****	
Giving or accepting bribes, kickbacks, or inappropriate gifts	**	
Falsifying records and reports	****	
Lying to employees, customers, vendors, or the public	****	
Withholding needed information from customers, vendors, or the public	***	
Misreporting actual time or hours worked	*	
Stealing, theft, or related fraud		
Breaking environmental and safety laws or regulations	****	***
Abusing or intimidating other company employees	****	
Discriminating on the basis of race, color, gender, age, or similar categories	*	
By Vendor Employees:		
Abusing drugs or alcohol		
Engaging in sexual harassment	***	
Giving or accepting bribes, kickbacks, or inappropriate gifts	**	
Falsifying records and reports	*	
Lying to employees, customers, vendors, or the public		
Withholding needed information from customers, vendors, or the public		
Misreporting actual time or hours worked		
Stealing, theft, or related fraud	*	
Breaking environmental and safety laws or regulations		
Abusing or intimidating other company employees		
Discriminating on the basis of race, color, gender, age, or similar categories		*
By Customer Employees:		
Abusing drugs or alcohol		
Engaging in sexual harassment	**	
Giving or accepting bribes, kickbacks, or inappropriate gifts		
Falsifying records and reports	*	*

*p=.05
**p=.01
***p=.001
****p=.000

DISCUSSION

Overall, the findings of this study showed dramatic differences between the responses of LEs and SMEs. Starting with the existence of written codes of ethics, we found codes to be common in LEs and far less common in SMEs. In addition, codes were communicated more formally to employees in LEs than in SMEs and in LEs mandatory training was used to assure that employees were knowledgeable about the corporate codes and values. One possible explanation for these differences is that in SMEs, there is less formality in the workplace, leading to more frequent interactions between employees and management. This type of communication and interaction would reduce the necessity for formal written codes. In contrast, in large firms, employees are more removed from management and must rely on written documents to learn about corporate values. These conclusions are similar to those drawn by Spence and Lozano [69], Vyakarnam, et al [85], Perrini et al [53], Graafland [27] in their work comparing SMEs and LEs.

Since SMEs indicated that written codes of ethics are not critical to their awareness of corporate values, it follows that neither the written codes nor mandatory training be needed to demonstrate top management or employee commitment to social responsibility. In contrast, the mere existence of formal codes in LEs influenced perceptions that employees and top management are committed to social responsibility. In fact, the prevalence of mandatory training in LEs strengthened these perceptions. The impacts that the code and training had on perceptions of commitment in LEs makes sense if one considers that both of these undertakings represent a significant investment on the part of the LEs, thereby signaling to employees the importance that management places on ethics and social responsibility. These findings are consistent with the finding of Graafland, et al [27] and Spence and Lozano [70]. Since mandatory training did strengthen perceptions of employees in LEs, it can be argued that perceptions about employee and top management commitment to social responsibility in SMEs could be strengthened if mandatory training was utilized.

An interesting result from our study was the fact that both LEs and SMEs expect vendors and B2B customers to behave ethically even when no formal communication of these expectations had occurred. These results for both large and small firms seem to suggest a general belief that ethical behavior stems primarily from personal values and doesn't need to be dictated by codes.

With respect the whether codes of conduct have impacted behaviors in LEs or SMEs, as you can see from Tables 2, 3 and 4 mandatory training, top management commitment, and employee commitment had significant impacts on virtually every type of misconduct for LE's own employees. In contrast, for SMEs, only one type of misconduct (breaking environmental and safety laws or regulations) was affected by these three categories.

In terms of whether mandatory training, top management commitment, or employee commitment impacted the behaviors of vendor employees, we discovered that for LEs mandatory training has no impact on vendor employees behavior whereas employee and

top management commitment did have major impacts. For SMEs, only one type of misconduct (discriminating on the basis of race, color, gender, age, or similar categories) was impacted by mandatory training, top management commitment, or employee commitment.

REFERENCES

[1] Adams, J.S., Tashchian, A. and Shore, T. Codes of Ethics as Signals for Ethical Behavior. Journal of Business Ethics, 2001, 29, 199-211.

[2] Allen, J. and Davis, D. Assessing Some Determinant Effects of Ethical Consulting Behavior: The case of Personal and Professional Values, Journal of Business Ethics**,** 1993,12, 449-458.

[3] Amaeshi, K., Osoouji, O. and Nnodim, P. Corporate Social Responsibility in Supply Chains of Global Brands: A Boundaryless Responsibility? Journal of Business Ethics, 2008, 81,223-234.

[4] Arlow, P. and Ulrich, T. A longitudinal Survey of Business School Graduates' Assessments of Business Ethics, Journal of Business Ethics, 1988, 7(4), 295-302.

[5] Asgary, N. and Mirschow M.C. Toward a Model for International Business Ethics, Journal of Business Ethics, 2002, 36, 239-246.

[6] Badaracco, Jr., J.L. and Webb, A.P. Business Ethics: A View From the Trenches, California Management Review, 1995, 37 (2), 8-28.

[7] Baumhart, R. Problems in Review: How Ethical are Businessmen? Harvard Business Review, 1961, 39, 6-9.

[8] Biedermann, R. From a Weak Letter of Intent to Prevalence: The Toy Industries' Code of Conduct**,** Journal of Public Affairs, 2006, 6 (3/4) Aug/Nov., 197-209.

[9] Brief, A.P., Dukerich, J.M., Brown, P.R., and Brett, J.F. What's Wrong with the Treadway Commission report? Experimental Analyses of the Effects of Personal Values and Codes of Conduct on Fraudulent Financial Reporting, Journal of Business Ethics, 1996, 15, 183-198.

[10] Carasco, E.F. and Singh, J.B. The Content and Focus of the Codes of Ethics of the World's Largest Transnational Corporations, Business and Society Review, 2003, 108, 71-94.

[11] Carter, C.R. and Jennings, M.M. The Role of Purchasing in Corporate Social Responsibility: A Structural Equation Analysis, Journal of Business Logistics, 2002, 25 (1), 145-186.

[12] Center for Business Ethics, Instilling Ethical Values in Large Corporations, Journal of Business Ethics, 1992, 11, 863-867

[13] Clarkson, M.B.E. and Deck, M.C. Effective Codes of Ethics: A Stakeholder Approach, The Clarkson Centre for Business Ethics, Toronto, Ontario, 1992.

[14] Cooper, R.W., Frank, G.L. and R.A. Kemp, R.A. Ethical Issues, Helps and Challenges: Perceptions of Members of the Chartered Institute of Purchasing and Supply" European Journal of Purchasing and Supply Management, 1997, 3, 189-198.

[15] Crawford, K.M. and Cox, J.F. Addressing Manufacturing Problems through the Implementation of Just-In-Time, Production and Inventory Management Journal, 1991, 32 (1), 33-36.

[16] Cressey, D. R. and Moore, C.A. Managerial Values and Corporate Codes of Ethics, California Management Review, 1983, 25, 53-77.

[17] Detindamar, D. Corporate Social Responsibility Practices and Environmentally Responsible Behavior: The Case of The United Nations Global Compact, Journal of Business Ethics, 2007, 76, 163-176.

[18] Ebrahimpour, M. and Withers, B.E. A Comparision of Manugacturing Management in JIT and non-JIT Firms, International Journal of Production Economics, 1993, 32, 353-364.

[19] Egels-Zanden, N. Suppliers' Compliance with MNCs' Codes of Conduct: Behind the Scenes at Chinese Toy Suppliers, Journal of Business Ethics, 2007, 75, 45-62.

[20] Estes, R. *Tyranny of the Bottom Line*, Berrett-Koehler Publishers, San Francisco, CA 1996.

[21] Ethical Trading Initiative: 1998, http://www.triadcraft.co.uk./etil.htm.

[22] Ethics Research Center, Creating a workable Company Code of Ethics (Washington, DC, 1990.

[23] Ethics Research Center, Ethics in American Business: Policies, Programs, and Perceptions, Washington, DC, 1994.

[24] Ferrell, O.C. and Gresham, L. A Contingency Framework for Understanding Ethical Decision Making in Marketing, Journal of Marketing, 1985, 49, 87-96.

[25] Fransen, L. and Kolk, A. Global Rule-Setting for Business: A Critical Analysis of Multi-Stakeholder Standards, Organization, 2007, 14 ,667-684.

[26] Frenkel, S.J. and Scott, D. Compliance, Collaboration and Codes of Labor Practice, California Management Review , 2002, 45, 29-49.

[27] Graafland, J., van de Ven, B., and N. Stoffle, N. Strategies and Instruments for Organising CSR by Small and Large Businesses in the Netherlands, Journal of Business Ethic,*s* 2003, 47, 45-60.

[28] Grundstri-Amado, R.A Strategy for Formulation and Impementation of Codes of Ethics in Public Service Organizations, Journal of Public Administration,2001, 24, 461-478.

[29] Handfield, R. B. and Baumer, L. Managing Conflict of Interest Issues in Purchasing, Journal of Supply Chain Management: A Global Review of Purchasing and Supply ,2006, 42, 41-50.

[30] Harich, K. and Curren, M. A longitudinal Examination of American Business Ethics: Clark's Sacles Revisted, Business and Professional Ethics Journal ,1995, 14(4), 57 -68.

[31] Helin, S. and Sandström, J. An Inquiry into the Study of Corporate Codes of Ethics, Journal of Business Ethic,*s* 2007, 75, 253-271.

[32] Hemphill, T. and Lillevik, w. The Global Economic Ethic Manifesto: Implementing a Moral Values Foundation in the Multinational Enterprise, Journal of Business Ethics. 2011, **101**, 213-230.

[33] Hunt, S.D., and Vitell, S. A General Theory of Marketing Ethics, Journal of Macromarketing, 1986, 8, 5 – 16.

[34] Isaksson, R., P. Johansson, and Fischer, K. Detecting Supply Chain Innovation Potential for Sustainable Development, Journal of Business Ethics, 2010, 97, 425-552.

[35] Jiang, B. Implementing Supplier Codes of Conduct in Global Supply Chains: Process Explanations from Theoretic and Empirical Perspectives, Journal of Business Ethics, 2009, 85, 77-92.

[36] Kaptein, M. and Schwartz,M.S. Business Codes of Multinational Firms: What do They Say? Journal of Business Ethic,*s* 2008, 50, 13-31.

[37] Kolk, A. and van Tulder, R. Child Labor and Multinational Conduct: A Comparison of International Business and Stakeholder Codes, Journal of Business Ethics ,2002. 36, 291-306.

[38] Kaufman, A., Tiantubtim, E., Pussayapibul, N., and P. Davids, P. Implementing Voluntary Labour Standards and Codes of Conduct in the Thai Garment Industry, Journal of Corporate Citizenship, 2004, Issue 13 (Spring), 91-99.

[39] Ladd, J.The Quest for a Code of Professional Ethics, in D.G. Johnson and J.W. Snapper (eds.), Ethical Issues in the Use of Computers, Wadsworth, Belmont, CA, pp. 37-38, 1985.

[40] Leipziger, D.The Corporate Responsibility Code Book Greenleaf Publishing, Sheffield, 2003.

[41] Longenecker, J.G., Moore, C.W., Petty, J.W., Palich, L.E., and McKinney, J.A. Ethical Attitudes in Small Businesses and Large Corporations: Theory and Empirical Findings from a Tracking Study Spanning Three Decades, Journal of Small Business Management, 2006, 44(2), 167-183.

[43] Maloni, M.J. and Brown, M.E. Corporate Social Responsibility in the Supply Chain: An Application in the Food Industry, Journal of Business Ethics, 2006, 68, 35-52.

[44] Mamic, I. Managing Global Supply Chain: The Sports Footwear, Apparel and Retail Sectors,. Journal of Business Ethic*.s* 2005, 59(1/2), 81-100.

[45] Mathews, C.M. Codes of Ethics: Organizational Behavior and Misbehavior, in W.F. Frederick (ed.), Research in Corporate Social Performance and Policy Vol. 9, JAI Press, Greenwich, CT, pp. 10-130, 1987.

[46] Mazar, N., Amir, O., and Arierly, S. The Dishonesty of Honest People: A Theory of Self-Concept Maintenance, Journal of Marketing Research, 2008, XLV, 633-644.

[47] McKinney, J.A, Emerson, T.L., andNeubert, M.J. The Effects of Ethical Codes on Ethical Perceptions of Actions toward Stakeholders, Journal of Business Ethics, 2010, 97, 505-516.

[48] Molander, E.A. A Paradigm for Design, Promulgation and Enforcement of Ethical Codes, Journal of Business Ethics. 1987, 6(8), 619-631.

[49] Murphy, P.R., Smith,J.E., and Daley, J.M. Executive Attitudes, Organizational Size and Ethical Issues: Perspectives on a Server Industry, Journal of Business Ethic*,s* 1992, 11, 11-19.

[50] Ngai-Ling Sum and Ngai, P. Globalization and Paradoxes of Ethical Transnational Production: Code of Conduct in a Chinese Workplace, Competition & Change, 2005, 9(2), 181-200.

[51] O'Dwyer, B. and Maden, F. Ethical Codes of Conduct in Irish Companies: A Survey of Code Content and Enforcement Procedures, Journal of Business Ethics, 2006, 63, 217-236.

[52] O'Rourke, D. Multi-Stakeholder Regulation Privatizing or Socializing Global Labor Standards? World Development ,2006, 34(5), 899-918.

[53] Peretson, L.E. ans Krings,R. Are Ethical Codes of Conduct Toothleess Tigers for Dealing with Employment Discrimination?, Journal of Business Ethics, 2009, 85, 501-514.

[54] Perrini, F., Russo, A., and Tencati, A. CSR Strategies of SMEs and Large Firms, Evidence from Italy, Journal of Business Ethics, 2007, 74, 285-300.

[55] Piercy, N.F. and Lane, N. Ethical and Moral Dilemmas Associated with Strategic Relationships between Business–to-Business Buyers and Sellers, Journal of Business Ethics, 2007, 72, 87-102.

[56] Preuss, L. Rhetoric and Reality of Corporate Greening: A View from the Supply Chain Management Function" Business Strategy and the Environment, 2005, 14, 123-139.

[57] Pruess, L. Ethical Sourcing Codes of Large UK-Based Corporations: Prevalence, Content, Limitations, Journal of Business Ethics, 2009, 88, 735-747.

[58] Pruess, L. Codes of Conduct in Organisational Context: From Cascade to Lattice-Work of Codes, Journal of Business Ethics, 2010, 94, 471-487.

[59] Raj-Reichert, G. The Electronic Industry Code of Conduct: Private Governance in a Competitive and Contested Global Production Network, Competition & Change, 2011, 15(3), 221-238.

[60] Roberts, S. Supply Chain Specific? Understanding the Patchy Success of Ethical Sourcing Initiatives, Journal of Business Ethics, 2003, 44(2/3), 159-170.

[61] Russo, A. and Tencati, A. Formal vs. Informal CSR Strategies: Evidence from Italian Micro, Small, Medium-sized, and Large Firms, Journal of Business Ethics, 2009, 85, 339-353.

[62] Saini, A. Purchasing Ethics and Inter-Organizational Buyer-Supplier Relational Determinants: A Conceptual Framework, Journal of Business Ethics, 2010, 95, 439-455.

[63] Salam, M.A. Corporate Social Responsibility in Purchasing and Supply Chain, Journal of Business Ethics, 2009, 85, 355-370.

[64] Schlegelmilch, B.B. and Houston, J.E. Corporate Codes of Ethics in Large UK Companies: An Empirical Investigation of Use, Content and Attitude" European Journal of Marketing, 1989, 23,7-24.

[65] Schonberger, R.J., Japanese Manufacturing Techniques: Nine Hidden Lessons in Simplicity, Free Press, NY, 1982.

[66] Schwartz, M.S. The Nature of the Relationship Between Corporate Codes of Ethics and Behavior, Journal of Business Ethics, 2001, 32, 247-262.

[67] Schweitzer, M.E., Ordóñez, L., and Douma, B. Goal Setting as a Motivator of Unethical Behavior, Academy of Management Journal, 2004, 47, 422-432.

[68] Sethi, S., Veral, E., Shapiro, H., and Emelianova,O. Mattel, Inc.: Global Manufacturing Principles (GMP) - A Life-Cycle Analysis of a Company-Based Code of Conduct in the Toy Industry, Journal of Business Ethics, 2011, 99, 483-517.

[69] Smales, P. Living Wages and Institutional Supply Chain Duties, Business & Professional Ethics Journal, 2010, 29 (1-4), 109-134.

[70] Spence, L.J. and Lozano, J.F. Communicating About Ethics with Small Firms: Experiences from the U.K. and Spain, Journal of Business Ethics, 2000, 27(1), 43-53.

[71] Stevens, B. The Ethics of the US Business Executive: A Study of Perceptions, Journal of Business Ethic,*s* 2004, 54, 163-171.

[72] Stevens, B. Corporate Ethical Codes: Effective Instruments for Influencing Behavior, Journal of Business Ethics, 2008, 78, 601-609.

[73] Stevens, G. Business Ethics and Social Responsibility: The Repsponse of Present and Future Managers, Akron Business and Economics Review, 1984, 15, 6-11.

[74] Stevens, J. M., Steensma, J.K., Harrison, D.A., and Cochran,P.L. Symbolic or Substantive Document? The Influence of Ethics Codes on Financial Executives' Decisions, Strategic Management Journal, 2005, 26, 181-195.

[75] Stohs, J.H. and Brannick, T. Code and Conduct: Predictors of Irish Managers' Ethical Reasoning, Journal of Business Ethics, 1999, 22, 311-326.

[76] Strand, R. Corporate Responsibility in Scandinavian Supply Chain, Journal of Business Ethic,*s* 2009**,** 85, 179-185.

[77] Trevino, K.T. Ethical Decision Making in Organizations: A Person-Situation Interaction Model, Academy of Management Review ,1986, 11(3), 601-617.

[78] Trevino, L.K. and Weaver. G.R. Organizational Justice and Ethics Program 'Follow-Through': Influences on Employees' Harmful and Helpful Behavior, Business Ethics Quarterly, 2001, 11, 651-671.

[79] Tucker L.R., Stathakopolous, V., and C. Patti, C. A Multidimensional Assessment of Ethical Codes: The Professional Business Association Perspective, Journal of Business Ethics, 1999, 19, 287-300.

[80] Valentine, S. and Barnett, T. Ethics Codes and Sales Professionals' Perceptions of Their Organizations' Ethical Values, Journal of Business Ethic*.s* 2002, 40, 191-200.

[81] Van Tulder, R. and Kolk, A. Multinationality and Corporate Ethics: Codes of Conduct in the Sporting Goods Industry, Ans. Journal of International Business Studies, 2001, 32,2, 267-284.

[82] Veseth, M. Selling Globalization: the Myth of the Global Economy, London: Lynne Riener Publishers, 1998.

[83] Victor, B. and Cullen, J. The Organizational Bases of Ethical Work Climates, Administrative Science Quarterly, 1998, 33, 101-125.

[84] Vittell, S.J., Dickerson, E.B., and Festervand, T.A. Ethical Problems, Conflicts, and Beliefs of Small Business Professionals, Journal of Business Ethics, 2000. 28(1), 15-24.

[85] Von Weltzein, H. and Mele, D. Can an SME Become a Global Corporate Citizen? Evidence from a Case Study, Journal of Business Ethics, 2009, 88, 551-563.

[86] Vyakarnam, R., Bailey, A., Myers,A., and Burnett, D. Towards an Understanding of Ethical Behavior in Small Firms, Journal of Business Ethics, 1997, 16, 1625-1636.

[87] Weaver, G.R., Trevino, L.K., and P. Cochran, P. Corporate Ethics Practices in the Mi-1990's: An Empirical Study of the Fortune 1000, Journal of Business Ethics, 1999, 18, 283-294.

[88] Withers, B.E., Ebrahimpour, M., and Hikmet, N. An Exploration of the Impact of TQM and JIT on ISO 9000 Registered Companies, International Journal of Production Economics, 1997, 53, 209-216.

[89] Wood, G. and M. Rimmer, M. Codes of Ethics: What are They Really and What Should They Be? International Journal of Value-Based Management, 2003, 16, 181-195.

[90] Wotruba, T.R. A Comprehensive Framework for the Analysis of Ethical Behavior, with a Focus on Sales Organizations, Journal of Personal Selling and Sales Management, 1990, 10, 29-42.

[91] Yu, X. Impacts of Corporate Code of Conduct on Labor Standards: A Case Study of Reebok's Athletic Footwear Supplier Factor in China, Journal of Business Ethic,*s* 2008, 81, 513-529.

[92] Yu, X. From Passive Beneficiary to Active Stakeholder: Workers' Participation in CSR Movement Against Labor Abuses, Journal of Business Ethics, 2009, 87, 233-249.

GREEN OPERATIONS AND ORGANIZATIONAL PERFORMANCE

Thomas Ngniatedema, Kettering University, Flint, MI, (810)762-7956, tngniate@kettering.edu
Suhong Li, Bryant University, Smithfield, RI, (401) 232-6503, sli@bryant.edu

ABSTRACT

This study investigates the influence of green operations on organizational performance for top 500 publicly traded companies in the US. The results show that the companies in manufacturing industry have a lower score in environmental impact and a higher score in green reputation than those in services industry. In addition, impacts of green operations on manufacturing and services industries are different. In manufacturing industry, environmental impact score and green policies and performance score have an impact on organizational performance; while in services industry, green reputation plays a more important role in impacting its overall organizational performance.

Keywords: Green operations, environmental impact, green policies and performance, green Reputation

INTRODUCTION

Today, environmental protection and economic burden on industry is leading organizations toward reevaluation of their corporate strategies including operations and business performance. This is reflected in the growing number of recent papers exploring the relationship between environmental operations and business performance [2] [28] [32] [39]. Some empirical evidence suggesting a positive relationship between "green" operations and business performance is also cited in the literature [21]. However, more empirical work is needed to clarify the nature of this relationship [23] [26].

For the most part, research on environmental operations has focused on the areas of green product and process development, lean and green operations management, and remanufacturing and closed-loop supply chains [1] [28]. It can be argued that few studies in the area of operations management have looked at the relationship between green operations and firm performance in manufacturing and service industry simultaneously. Existing empirical evidence so far has been limited to manufacturing sector [21] [25] [27]. On the other hand, few studies have explored environmental issues in the service industry [15] [17] [21]. Very few of these studies found scientific evidence that green practices have important effect on firm performance. For instance, Enz and Siguaw [11] and Schendler [44] argue that environmental practices can improve customer loyalty and employee satisfaction, reduce costs, and enhance competitiveness. Reviewing published studies exploring the relationship between green operations and firm performance in the service industry, Kassinis and Soteriou [21] found most to be limited to the manufacturing industry and to be case studies, which predominantly identified opportunities for future research.

While numerous environmental frameworks, cases and concepts have evolved around

manufacturing industry, we are in the midst of a service revolution that is rapidly transforming industries and changing some fundamental assumptions we have about business and economics. Increasingly, the size of the service economy's contribution to gross domestic product is more than 70% in the U.S. and other developed countries, while the share of employment in services exceeds 80% in the U.S. and continue to rise [14] [41], which implies that we need to have a perspective and insight into these issues as they are becoming relevant to almost any organizational stakeholder. In addition, the pervasive environmental issues and their implications for business present tremendous opportunities to traverse growing and changing knowledge gaps regarding manufacturing (which focuses on creating and producing products and structures) and services (which focus on creating and delivering added value that is essentially intangible) [42]. This gap, long acknowledged by Kassinis and Soter [21], who wrote that "In practice, we know little about the environmental impacts of most service operations, how they can be managed, and what impact the environmental practices service firms adopt have on performance" prompts us to look into the relationship between green operations and firm's financial performance in both manufacturing and service industry simultaneously.

This paper identifies three key environmental operations that are important antecedents to a firm's financial performance in the context of manufacturing and services industry. In the next section, we review the literature about environmental operations and practices as well as firm-level performance, then, we raise our research questions and put forward a theoretical perspective to explain the relationship between these green operations and firm's performance. Then present our empirical investigation using data from Compustat, a database of financial, statistical and market information on active and inactive global companies throughout the world. and Newsweek, an information gatekeeper that enables consumers to access a list of environmental friendly companies. Following the methodology and the analysis, we interpret our findings and outline implications and future research.

THEORETICAL FOUNDATION AND HYPOTHESES DEVELOPMENT

Green Operations

Traditionally, environmental issues have attracted the attention of researchers in various areas of operations management. The scope of research ranges from studying operational problems such as green product and process development, lean and green operations management, to remanufacturing and closed-loop supply chains [28] [3] [6]. Environmental perspectives on operations lead to different terminologies with varying scope. One term emerging from the literature is "green operations". It relates to all aspects related to product manufacturing, usage, handling, logistics and waste management once the design has been finalized [47].

To gauge business processes excellence related to firm operations, increasingly, the research community is adding green metrics that allow key practitioners to monitor environmental efforts for themselves as well as to make informed decisions accordingly [16]. For example, Trucost, which specializes in quantitative environmental performance measurement; and CorporateRegister.com, the world's largest online directory of social responsibility, sustainability and environmental reporting used the terms "environmental impact score", "green policies and performance score", and "reputation survey score" to assess environmentally responsible (green)

practices among firms. According to these sources, a company's environmental impact score is overall environmental impact of a company's global operations. Green policies and performance score are the rules and guidelines that regulate all operations of a company. Finally, green reputation reflects the public image of a company in relation to its attitude and actions toward environmental issues when managing its operations and product lines. A number of authors have proposed research frameworks to assess business performance of environmental responsible firms. . Beamon [4] describes performance measures appropriate for the extended supply chain. Labuschagne and van Erck [31] and Chinander [5] propose a framework and methods for incorporating environmental objectives into firm's operations and performance. Building on these environmental score and concepts from Trucost and CorporateRegister.com, we study the relationship between green operations (measured by environmental impacts, green policies and performance and reputation survey) and firm performance in manufacturing and service industry.

Organizational Performance

Research conducted on environmental issues over the last decade use numerous approaches to measure firm financial performance. Financial measures such as return on sales (ROS), return on assets (ROA), return on equity (ROE), and return on invested capital (ROIC) have been used in the literature to evaluate the interests of various stakeholders in the market place [18] [19] [29] [43]. To model capital borrowed by stockholders from creditors and investors as well as their equity capital contribution, authors such as Konar and Cohen [30], Russo and Fouts [40], Elsayed and Paton [12], Nakao et al. [35], and King and Lenox [22 use ROA, Tobin's $q-1$, ROS, ROE, and return on capital employed (ROCE) to measure firm financial performance; where Tobin's $q-1$ was interpreted as the intangible assets value of the firms. Using the argument that managers are more open to offering their perceptions rather than offering precise quantitative data, other scholars use subjective perceptions of managers to assess firm financial performance [7] [20] [45].

This study is concerned with firm's performance relative to the market and its competition. Debt ratio (DR), profit margin (PM), return on total assets (RTA), and market to book ratio (MBR) are recognized as important dimensions to firm's financial performance [46]. DR is defined as the total debt over total assets. PM, a primary variable most investors examine when analyzing a company's performance, measures the profitability of a company and represents the net income over the sales. The RTA represents the net income over the total assets, and the MBR represents the market price over the book value.

Relationship between Environmental Practices and Organizational Performance

In studying environmental issues, the extant literature finds a significant correlation between green practices and corporate profitability within any organization. It is reported that companies having higher score according to some objective environmental criteria realize stronger financial returns than the overall market, whereas companies with poor score have weaker returns [7] [13] [32] [34].The prevailing view is that incorporating environmental variables into firms' activities often impacts costs because additional requirements have to be met to this end. This in turn impacts firm-level financial performance [36] [38].

Empirical studies that have analyzed the relationship between environmental operations and practices and firm's financial performance are fragmented across industries. Widely-cited research results relate environmental operations and practices to stock market performance, firm's market valuation, and competitive advantage [6] [19] [29]. Most of these studies suggest that environmental performance is positively correlated with the intangible asset value of S&P 500 firms as well as firm market value [10] [26] [22] [30]. A study by Nakao et al. [35] reveals that for the particular case of Japanese manufacturing sector, environmental performance improves ROA and Tobin's $q - 1$. From a competitive perspective, Porter and van der Linde [36], Rao and Holt [37], Dao et al. [8], and Reinhardt [38] suggest that environmental operations can improve firm-level financial performance and overall competitiveness through green products or services. The same authors recognize that poor environmental performance can reduce a firm's market valuation. The literature also reveals "green" firms to be more efficient and innovative [23] [36].

Although most studies find a positive correlation between environmental performance and firm-level performance, some results are conflicting and ambiguous [22]. For example, Kiernan [24] and Derwall et al. [9] show that environmental performance and firm-financial performance is negatively correlated. A study by Min and Galle [33] suggests that compared to liabilities and product disposal costs, competitive advantage plays a relatively minor role for managers considering green purchasing. In addition, Walley and Whitehead [49] argue that corporate environmental initiatives generate unrecoverable costs, divert resources from other productive investments, and conclude that they are unsustainable.

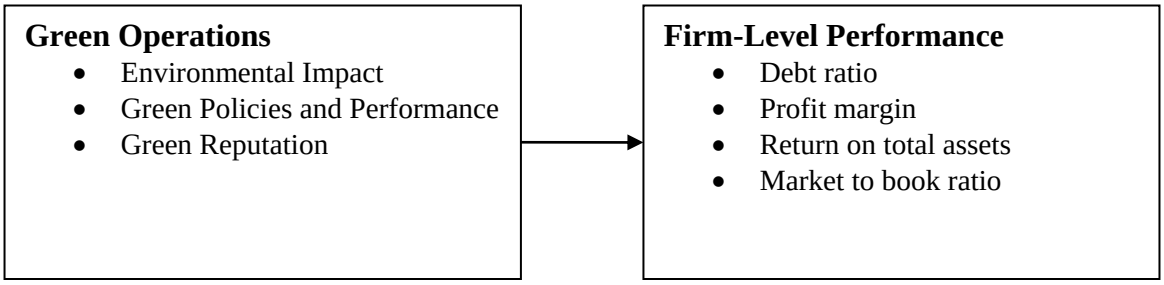
While previous approaches linking environmental operations and practices to firm-level performance seem to be fairly comprehensive and contribute significantly to our knowledge, existing empirical evidence so far has been limited to manufacturing sector [25] [27]. Environmental operations and practices, however, have been shown to be an important component of a service firm's operations [21]. Although a handful number of research papers study environmental issues in the context of services industry in general, Kassinis and Soteriou [21] argue that the results reported in these papers, however, is "limited by the case study or anecdotal nature of the evidence they are based on" and acknowledge the need for further empirical work to assess the relationship between environmental practices and firm-level performance in the service industry. The relative scarcity of research examining these two variables and the fact that prior approaches do not report consistent findings with regard to explaining how green operations impact firm-level performance both in manufacturing and service industry [10] [22] [26] motivates the need for this study. It is our contention that by jointly examining the impact of environmental practices to firm-level performance in manufacturing and service industry, we can substantially contribute toward the findings of earlier studies.

Research Framework

Figure 1 schematically represents our research framework. We draw on concepts from interrelated literature streams of environmental operations, practices and corporate growth to propose a research model that assesses a direct effect between green operations and firm-level performance. Our framework suggests that firm-level performance is impacted by three green

operation factors: environmental impact, green policies and performance, and green reputation. Other potential factors may impact firm-level performance and not included in his study due to the limitation of the data.

Figure 1: Research Framework



RESEARCH METHODOLOGY

The focus of this research is top 500 publicly traded companies by revenue, market capitalization and number of employees in 2010. The score of green operations for each company was obtained from Newsweek. The financial performance of each company was obtained from Compustat. In addition, each company was classified into manufacturing or services industry based on its major business activities.

Data Analysis and Discussion

This section will first analyze whether there is a significant difference on overall green operations between Manufacturing and Service industry using t-tests. Second, six regression analysis will be conducted to see which green indictors have an impact on each financial performance (Debt Ratio, Profit Margin, Return on Total Assets, and Market to Book Ratio) in manufacturing and services industry respectively.

Green Operations between Manufacturing and Services Industry

A series of t-tests were conducted to see whether there is a difference in overall score and the score for each green operations measure. The results are shown in Table 1. It can be seen that there exist significant difference in environmental impact and reputation survey score between manufacturing and services industry. The companies in manufacturing industry have a lower score in environmental impact and a higher score in reputation survey than those in services industry. No significant differences were found in term of green policies and performance score in both industries.

Green Operations and Financial Performance in Manufacturing Industry

Table 2 shows the results of green operations on each indicator of financial performance in manufacturing industry. It can be seen that green operations have a significant impact on three of

four financial indicators (Debt Ratio, Profit Margin and Market to Book Ratio). Specially, a higher environmental impact score is negatively associated with debt ratio and is positively associated with profit margin, indicating that a company's initiatives to reduce environmental impact of its operation does lead to a low debt ratio and a high profit margin. The findings also show that a higher score in green policies and performance leads to a higher market-to-book ratio, indicating positive stock market reaction to green initiative.

Table 1 t-tests on Green Operations between Manufacturing and Service Industry

Green Operations Indicator	Industry Sector	N	Mean	t	Significance
Environmental Impact Score	1 (Manufacturing)	298	.42	-8.14	**0.00**
	2 (Services)	183	.63		
Green Policies and Performance Score	1 (Manufacturing)	298	.43	1.32	0.19
	2 (Services)	183	.41		
Reputation Survey Score	1 (Manufacturing)	298	.50	4.35	**0.00**
	2 (Services)	183	.44		

Green Operations and Financial Performance in Services Industry

Table 3 shows the results of green operations on each indicator of financial performance in services industry. The results show that green reputation score is associated positively with debt ratio and market to book ratio, indicating that a company with a good reputation in managing environmental impact has a higher debt ratio and a higher market-to book ratio.

Comparing the results of service industry with manufacturing industry, it can be seen that the impact of green operations on both industry are different. In manufacturing industry, environmental impact score and green policies and performance score will have an impact on a firm's financial performance; while in service industry, green reputation plays a more important role in impacting its financial performance.

CONCLUSION AND IMPLICATION

This study investigates the relationships between green operations and organizational performance for top 500 publicly traded companies in the US. Green operations were measured by three indicators (environmental impact, green policies and performance, and green reputation) and the scores of those indicators for each top 500 company in 2010 were obtained from Newsweek. Organizational performance were measured by debt ratio, profit margin, return on total assets and market to book ratio and were obtained from Compustat, a database of financial, statistical and market information on active and inactive global companies throughout the world. The results show that the companies in manufacturing industry have a lower score in environmental impact and a higher score in green reputation than those in services industry. In addition, impacts of green operations on manufacturing and services industries are different. In manufacturing industry, environmental impact score and green policies and performance score have an impact on organizational performance; while in services industry, green reputation plays a more important role in impacting its overall performance.

Table 2. Regression Analysis (Manufacturing Industry)

Independent Variables	Dependent Variables							
	Debt Ratio		Profit Margin		Return on Total Assets		Market to Book Ratio	
	Standardized Coefficient	Sig.	Standardized Coefficient	Sig.	Standardized Coefficient	Sig.	Standardized Coefficient	Sig.
Environment Impact Score	-.290	**.000**	.115	**.056**	-.062	.305	.008	.892
Green Policies and Performance Score	.104	.111	.104	.120	.112	.096	.168	**.012**
Reputation Score	.020	.756	.085	.202	.060	.367	.019	.771
R	.28		.17		.15		.18	
R^2	.08		.03		.02		.03	
F-statistics	8.42		2.94		2.12		3.32	
Significance	**.00**		**.03**		.09		**.02**	

Table 3. Regression Analysis (Service Industry)

Independent Variables	Dependent Variables							
	Debt Ratio		Profit Margin		Return on Total Assets		Market to Book Ratio	
	Standardized Coefficient	Sig.	Standardized Coefficients	Sig.	Standardized Coefficient	Sig.	Standardized Coefficient	Sig.
Environment Impact Score	-.125	.091	.018	.817	-.103	.165	-.138	.062
Green Policies and Performance Score	-.120	.129	-.007	.934	-.040	.615	.035	.653
Reputation Score	.168	**.033**	-.065	.420	-.135	.087	.175	**.026**
R	.22		.07		.19		.23	
R^2	.05		.01		.04		.05	
F-statistics	2.98		.29		2.22		3.37	
Significance	**.03**		.34		.09		**.02**	

REFERENCES

[1] Angell, L. C. and Klassen R. D. (1999), "Integrating Environmental Issues into the Mainstream: An Agenda for Research in Operations Management," *Journal of Operations Management*, 17, 5, 575–598.

[2] Aragon-Correa, J. A (1998), "Strategic Proactivity and Firm Approach to the Natural Environment," *Academy of Management Journal*, 41, 5, 556–567.

[3] Bai C., Sarkis J. Green supplier development: analytical evaluation using rough set theory. Journal of Cleaner Production (18) 2010, 1200-1210

[4] Beamon B.M., "Designing the Green Supply Chain". Logistics Information Management Volume 12. Number 4. 1999 . pp. 332±342.

[5] Chinander, K. R. 2001. Aligning accountability and awareness for environmental performance in operations. *Production and Operations Management* **10**(3) 276–291.

[6] Corbett, C. J. and Klassen R.D. (2006). Extending the Horizons: Environmental Excellence as Key to Improving Operations. Manufacturing & Service Operations Management 8(1), pp. 5–22.

[7] Correa A., Hurtado-Torres J., Sharma N., Garcia-Morales S.V.J., (2008). Environmental strategy and performance in small firms: a resource-based perspective. Journal of Environmental Management 86, 88–103.

[8] Dao V., Langella I., carbo J., "From green to sustainability: Information Technology and an integrated sustainability framework." Journal of Strategic Information Systems 20 (2011) 3–79.

[9] Derwall, J., N. Guenster, R. Bauer, K. Koedijk. 2005. The ecoefficiency premium puzzle. *Financial Analysts J.* 61(2) 51–63.

[10] Dowell, Glen; Hart, Stuart and Yeung, Bernard "Do Corporate Global Environmental Standards Create or Destroy Market Value?" *Management Science*, 2000 46(8): 1059-74.

[11] Enz, C. A. and Siguaw J. A. (1999), "Best Hotel Environmental Practices," *Cornell Hotel and Restaurant Administration Quarterly*, 40, 5, 72–77.

[12] Elsayed, K., Paton, D., 2005. The impact of environmental performance on firm performance: static and dynamic panel data evidence. Structural Change and Economic Dynamics 16 (3), 395-412.

[13] Estampe D., Lamouri S., Paris J-L., and Brahim-Djelloul S., "A framework for analysing supply chain performance evaluation models". Int. J. Production Economics, November 2010.

[14] Fitzsimmons, J. A. and Fitzsimmons M. J. (2000), *Service Management: Operations, Strategy and Information Technology*, 3rd ed., Irwin/McGraw-Hill, Boston.

[15] Foster, S. T., Sampson S. E., and Dunn S. C. (2000), "The Impact of Customer Contact on Environmental Initiatives for Service Firms," *International Journal of Operations and Production Management*, 20, 2,187–203.

[16] Golicic, S. L., Boerstler, C. N. and L. M. Ellram (2010). 'Greening' Transportation in the Supply Chain, *Sloan Management Review*, 51(2), 47-55.

[17] Goodman, A. (2000), "Implementing Sustainability in Service Operations in Scandic Hotels," *Interfaces*, 30, 3,202–214.

[18] Hart, S. L., Ahuja, G., 1996. Does it pay to be green? an empirical examination of the relationship between emission reduction and firm performance. Business Strategy and the Environment 5 (1), 30-37.

[19] Hiroki I. and Keisuke O., "How does environmental performance affect financial performance? Evidence from Japanese manufacturing firms." MPRA, 28. December 2010.
[20] Judge, W., Douglas, T., (1998). Performance implications of incorporating natural environmental issues into the strategic planning process: an empirical assessment. Journal of Management Studies 35, 241–262.
[21] Kassinis G.I. and Soteriou A.C. "Greening the Service Profit Chain: The Impact of Environmental Management Practices." Production and Operations Management, 12(3), 2003.
[22] King, A., Lenox, M., 2002. Exploring the locus of profitable pollution reduction. Management Science 48 (2), 289-299.
[23] King, A. A. and Lenox M. J. (2001), "Does it Really Pay to be Green? An Empirical Study of Firm Environmental and Financial Performance," *Journal of Industrial Ecology*, 5, 1, 105–116.
[24] Kiernan, M. J. 2001. Eco-value, sustainability, and shareholder value: Driving environmental performance to the bottom line. *Environ. Quality Management* 10(4) 1–12.
[25] Klassen, R. D. (1993), "The Integration of Environmental Issues into Manufacturing," *Production and Inventory Management Journal*, 34, 1, 82–88.
[26] Klassen, R. D., C. P. McLaughlin. 1996. The impact of environmental management on firm performance. *Management Sci.* 42(8) 1199–1214.
[27] Klassen, R. D. and Whybark D. C. (1999), "Environmental Management in Operations: The Selection of Environmental Technologies," *Decision Sciences*, 30, 3, 601–631.
[28] Kleindorfer, Singhal, and Van Wassenhove: **"***Sustainable Operations Management".* Production and Operations Management 14(4), pp. 482–492, 2005.
[29] Klingenberg B., and Geurts T.G., "A Theoretical Framework for Financial Performance Measurement of Operations Management Strategies". 2009 Northeast Decision Sciences Institute Proceedings.
[30] Konar, S., Cohen, M. A., 2001. Does the market value environmental performance? Review of Economics and Statistics 83 (2), 281-289.
[31] Labuschagne C.,Brent A.C., van Erck R. P. G., "Assessing the sustainability performances of industries", Journal of Cleaner Production 13 (2005) 373-385
[32] McCrea B. "Why "Green" Equals Good Business" Supply Chain Management Review, 2010
[33] Min, H., W. P. Galle. 1997. Green purchasing strategies: Trends and implications. *Internat. J. Purchasing Materials Management* 33(3) 10–17.
[34] Murphy M.J., The Profitable Correlation Between Environmental and Financial Performance: A Review of the Research, *LGA,(2002)*
[35] Nakao, Y., Amano, A., Matsumura, K., Genba, K., Nakano, M., 2007. Relationship between environmental performance and financial performance: an empirical analysis of Japanese corporations. Business Strategy and the Environment 16 (2), 106-118.
[36] Porter M.E. and van der Linde C., "Green and Competitive: Ending the Stalemate." Journal of Business Administration and Poli... Annual 1999 p 215, HBR, Sept.-Oct. 1995.
[37] Rao P., Holt D., "Do green supply chains lead to competitiveness and economic performance?" International Journal of Operations & Production Management Vol. 25 No. 9, 2005

[38] Reinhardt, F. L. (1999). Bringing the environment down to earth. Harvard Business Review 77(4) 149–179.
[39] Rothenberg, S., Pil F. K., Maxwell J.. 2001. "Lean, green, and the quest for superior performance." Production and Operations Management **10**(3) 228–243.
[40] Russo, M. V., Fouts, P. A., 1997. A resource-based perspective on corporate environmental performance and profitability. Academy of Management Journal 40 (3), 534-559.
[41] Salzman, J. (2000), "Environmental Protection Beyond the Smokestack: Addressing the Impact of the Service Economy," *Corporate Environmental Strategy*, 7, 1, 20–37.
[42] Sasser, W. E., Olsen R. P., and Wyckoff D. D. (1978), *Management of Service Operations*, Allyn and Bacon, Boston, Massachusetts.
[43] Sarkis, J., Cordeiro, J. J., 2001. An empirical evaluation of environmental efficiencies and firm performance: pollution prevention versus end-of-pipe practice. European Journal of Operational Research 135 (1), 102-113.
[44] Schendler, A. (2001), "Trouble in Paradise: The Rough Road to Sustainability in Aspen," *Corporate Environmental Strategy*, 8, 4, 293–299.
[45] Sharma, S., Vredenburg, H., (1998). Proactive corporate environmental strategy and the development of competitively valuable organizational capabilities. Strategic Management Journal 19, 729–753.
[46] Slywotzky, A. J., Morrison, D. J., and Weber, K. *How Digital Is Your Business?,* Crown Business, New York, 2000.
[47] Srivastava, S.K., 2007. Green supply-chain management: a state-of-the-art literature review. International Journal of Management Reviews 9 (1), 53–80.
[48] Tutore I. " Key Drivers of Dorporate Green Strategy", EDAMBA Summer Academy 2010
[49] Walley, N. and B. Whitehead (1994), "It's Not Easy Being Green," *Harvard Business Review*, 72, 3, 46–52.

The Role of Solar Energy in Sustainable Practices in Manufacturing Companies

Said Dini, PhD
Mechanical Engineering Department
Western New England University
Springfield, MA 01119

Richard B. Mindek, Jr., PhD
Mechanical Engineering Department
Western New England University
Springfield, MA 01119

ABSTRACT

In recent years, issues of sustainability have received much greater attention from organizations large and small, public and private. The level of environmental consciousness on the part of the average person has been dramatically impacted: "going green" and "carbon footprint," along with other environmentally-themed phrases, have become household terms in a short span of time. To this end, many institutions of higher education have striven to be leaders in a new call for sustainability and environmental conservation. Western New England University is one these institutions that is proving itself a leader in sustainability.

To support our leadership role in sustainability, alternative energy laboratory experiences were developed to foster the Green Concentration in the mechanical engineering program at Western New England University. These laboratories, which give students hands-on experience and a better understanding of basic concepts in wind energy, solar energy, and fuel cell technology, utilize an Alternative Energy Active Learning Platform. Furthermore, the program focuses on understanding the crucial role that solar energy and other renewable sources can play in enhancement of sustainable and efficient operations in manufacturing companies.

This paper discusses the implications of the usage of alternative energy sources for manufacturing facilities. In addition, the study examines the impact of designing energy independent systems in a manufacturing environment through conversion to solar power and other renewable energy sources. The cost effectiveness and benefits of such conversion are also examined and discussed.

MODELING COMPREHENSIVE SUSTAINABLE VALUE IN TECHNOLOGY RESEARCH AND DEVELOPMENT DECISIONS

Daniel R. Ball
Department of Management & Decision Sciences
Leon Hess Business School
Monmouth University
400 Cedar Avenue
West Long Branch, NJ 07764-1898, USA
dball@monmouth.edu
001 (732) 923-4642

ABSTRACT

The current economic climate has led to an increased awareness of the importance of strategic decisions that help organizations achieve long-term sustainability. A broader definition of business sustainability is proposed that includes technological innovation research and development, improved supportive operations management processes, and domestic and global business success, in addition to the more traditional areas of environmental, social, and economic sustainability. This paper presents a theoretical framework that integrates each factor of sustainability into a conceptual, endogenously-driven model that can assist organizations when evaluating the potential for technology research and development investment decisions to promote long-term value creation across these multiple dimensions of sustainability.

Keywords
Sustainability, technology management, innovation, operations management, strategy

1 INTRODUCTION

The current domestic and global economic climate has brought to the forefront both the importance and challenges of making strategic technology research and development (R&D) investment decisions that promote sustainability for an organization. As the degree of global competitiveness has continued to increase, technology life cycles have been drastically reduced in many industries and, consequently, the need for careful usage of R&D resources has become paramount for the successful development of next-generation technologies. Although the general concept of sustainability has become an important part of society and business, actual definitions of sustainability have not always been clear and the development of effective strategic decision-making models is complex. To further confuse this issue, the growing complexity of global business networks has provided a need for a more comprehensive decision-making model to assist organizations in achieving and sustaining long-term value creation

through technology R&D strategies, including both innovation and the supporting operations management processes.

One of the pioneering efforts in the area of sustainability studied the potential issues resulting from an exponentially growing world population when compared with only linear expected growths of necessary resources based on technological advances [1]. This study raised doubts as to whether there would be sufficient resources to sustain the expected growth in global population. Gro Harlem Brundtland, the former Prime Minister of Norway, more formally defined sustainability as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" [2] [3]. Additional formal definitions of both *sustainability* and *sustainable development* were summarized in [4]. Here, sustainability was defined as: "…the capacity of a system to maintain output at a level approximately equal to or greater than its historical average, with the approximation determined by the historical level of variability" [5]; "…maximizing the net benefits of economic development, subject to maintaining the services and quality of natural resources over time" [6]; and "…sustainability of natural ecosystems can be defined as the dynamic equilibrium between natural inputs and outputs, modified by external events such as climatic change and natural disasters" [7]. Definitions of sustainable development include: "…development that meets the needs of current generations without compromising the ability of future generations to meet their needs and aspirations" [8]; and "…development that improves the quality of human life while living within the carrying capacity of supporting ecosystems" [9].

These definitions of sustainability and sustainable development stress the importance of considering both environmental and social concerns when making business decisions. For decisions pertaining to technology R&D, in addition to the supporting operational processes, there may be considerable confusion regarding how to balance the strategic concerns of organizational stakeholders within these sustainability constraints. In this context, it is important for businesses to consider a more expanded concept of *corporate sustainability* that more explicitly incorporates both environmental and social issues, but also includes a strategic management focus aimed at ensuring long-term organizational value creation. The concept of corporate sustainability was presented to include social, environmental, and financial criteria, with the following nine performance principles [10]: ethics; governance; transparency; business relationships; financial return; community involvement / economic development; value of products and services; employment practices; and the protection of the environment [11]. A very succinct view of a sustainable organization is one that is able to thrive in perpetuity while considering four coequal components: social, economic (i.e., profitable operations), environmental, and cultural [3]. These views of corporate sustainability may be used to guide strategic technology R&D investment decision-making into directions that promote long-term business, economic, social, and ecological value creation. The model presented in this paper is intended to capture each element of sustainability, include any endogenous impacts of interlinking factors and effects, and assist organizational decision-making across multiple categories of sustainable value creation.

2 ORGANIZATIONAL MOTIVATIONS FOR CORPORATE SUSTAINABILITY

Before any strategic R&D investment decision-making model is constructed, the motivations behind organizations striving to achieve corporate sustainability should be identified. One possible motivation may be for the organization to become publicly recognized as promoting corporate social responsibility (CSR) and achieve related marketing benefits. For example, the carpet fiber maker Anton experienced an increase of $4 million in sales when it became certified by Scientific Certification Systems as an "Environmentally Preferable Product" [12]. Another motivator may be to recognize cost savings due to waste reduction. For example, 3M's 3P program (i.e., "Pollution Prevention Pays") has strived to reduce emissions, solid waste, and energy and water usage; a result of this program has been significant reductions of pollutants and $11 million in cost savings [12].

Whereas these CSR marketing and waste reduction examples exhibited clear benefits across each criteria of corporate sustainability, strategic investments in other areas such as new technology innovation R&D and operational processes are not as clear. For example, whereas the investment in a new technology such as improved hybrid electric vehicles may be made to foster environmental sustainability, the resulting risks of either cannibalizing the current gasoline-powered product line or not yielding a successful new product may actually hurt profits and long-term value creation potential for a company, as well as hinder its position to be able to generate value continuously into the future. Furthermore, the increased electricity demands, as well as the production and disposal of more complicated battery components, may yield tradeoffs in the area of environmental sustainability. Therefore, the potential costs and benefits of these ventures are not as clear, thus making any investment in this area risky for an organization. However, there are also significant risks of succumbing to incumbent inertia and missing out on the future market opportunities if these investments are indeed successful. A comprehensive framework that can effectively model the complexities of these factors is necessary to assist with organizational R&D investment decision-making.

Regardless of the organization's motivation, generating sustainable corporate value is becoming an important strategic consideration for both domestic and global business environments. For example, organizations in the United States have suffered as past managerial decisions have led to a current economic and industrial climate that has proven to be unsustainable (e.g., housing market crashes; financial market instability; off-shoring of both manufacturing and service related jobs). These domestic issues within the United States have also made it difficult for firms to compete in the global marketplace. One area that helps govern a firm's global business decisions has been the development of various internationally-recognized standards. Examples from the International Organization for Standardization (ISO) include the following series: 9000 (Quality Management), 14000 (Environmental Management), and 26000 (Social Responsibility). These standards assist organizations in making decisions and following operating practices that are recognized internationally and promote the concept of corporate sustainability.

In summary, the current domestic and global business domains have created a situation where firms need to generate sustainable value, while also adhering to environmental, resource, and future operational and innovative constraints. It is imperative for firms to effectively utilize their resources to invest in next-generation technology investment projects that promote sustainability across a broad spectrum of areas. This paper considers the importance of developing models that assist with organizations strategically managing their R&D decisions to achieve full sustainability: business (e.g., technology, innovation, and operations management), environmental, social, and economic sustainability.

3 MEASURES OF SUSTAINABILITY

Organizations may possess a multitude of strategic objectives when considering decisions that may affect each of the areas of corporate sustainability. As previously discussed, the lack of clarity and general consensus surrounding the various definitions of sustainability may lead to confusion as to how organizations should actually make these decisions. Furthermore, the development of an accurate strategic decision-making model to assist with technology R&D decisions is especially difficult due to many factors including the identification of affected parameters, possible parameter correlations, projected impacts, and uncertain R&D results, future technology performance, and market conditions. A number of metrics have been proposed in an effort to assist with measuring sustainability.

The IPAT model was one of the first metrics proposed [13]. This model attempts to determine the environmental impact (I) due to human activity as a function of the population size (P), per capital affluence (A), and the technology involved in supporting each unit of consumption (T). Two other more recent indices include the Environmental Sustainability Index (ESI) and the Environmental Performance Index (EPI) and were created as an Initiative of the Yale Center for Environmental Law and Policy (YCELP) and the Center for International Earth Science Information Network (CIESIN) of Columbia University, in collaboration with the World Economic Forum and the Joint Research Centre of the European Commission. The ESI is calculated on a nation-state basis and factors the following areas: environmental systems, reducing environmental stresses, reducing human vulnerability to environmental stresses, societal and institutional capacity to respond to environmental challenges, and global stewardship [4] [14]. The EPI evaluates a country's environmental performance and considers the following two objectives: reducing environmental stresses on human health, and promoting ecosystem vitality and sound natural resource management [15]. Although the IPAT, ESI, and EPI are just a few examples of sustainability metrics, it should be noted that the focus of each metric is in the areas of human factors, and environmental and natural resource sustainability. The reader is referred to [4] for summaries of additional sustainability performance metrics.

Because technology R&D investment decisions also require the inclusion of additional factors such as potential R&D success and the related economic factors necessary to achieve corporate

sustainability, it is necessary to account for these factors and the actual costs associated with any decision with the goal of achieving this broader concept of sustainability. The *Triple Bottom Line* (TBL) accounting approach refers to the importance of measuring social (i.e., people), economic (i.e., profits), and environmental (i.e., planet) factors [16]. The importance of including external environmental costs (e.g., pollution costs) in addition to internal costs in a full cost accounting approach to measure corporate sustainability was stressed by [17]. The TBL accounting approach was extended and a sustainable bottom line was created by also integrating cultural factors [3]. A key consideration of these metrics is that business practices would not be sustainable unless environmental, social, and cultural factors were also considered. These factors may also be included in the sustainability model presented in this paper.

These previous metrics presented have either focused primarily on environmental and social welfare factors, or have viewed the economic aspects of sustainability as external "burden costs" to a firm (e.g., pollution costs). An alternative measure of corporate sustainability was proposed and termed *sustainable value* [18]. This sustainable value metric compares the net value added by a firm to the cost of utilized sustainability capital and evaluates a sustainability efficiency value. This sustainability efficiency value may then be compared to a benchmark to determine how efficient a given firm is at utilizing the sustainable capital and generating sustainable value. An important feature of this methodology is that it measures corporate sustainability as value-based, whereas the other metrics view environmental and social factors in terms of their negative impacts (i.e., "burden costs").

4 COMPREHENSIVE SUSTAINABLE VALUE CREATION

The sustainable value creation philosophy posed by [18] should be extended to firms seeking to make decisions pertaining to technology R&D investments and the supporting operations management processes. A model is therefore required that may be used to evaluate the potential for any technology R&D investment decision to generate both short- and long-term domestic and global business sustainability potential, while also promoting broader environmental, social, and economic sustainability. The modeling approach proposed in this paper specifically identifies a theoretical framework for approaching these types of strategic decisions, while incorporating the complex system dynamics. This framework also extends previous research by [19] that presented some general concerns of domestic and global sustainability, and more broadly defines sustainability to account for individual business/industrial sustainability challenges in light of additional business constraints such as government environmental regulations and other more traditional sustainability factors. For example, some factors that promote environmental sustainability may restrict business/industrial sustainability, and vice versa. Consequently, the comprehensive strategic model proposed in this paper incorporates multiple sustainability factors that link organizational success, technology/innovation developments, operations management, and ecological/environmental, social, and economic sustainability.

The first step when developing an appropriate model is to identify each factor that may have an impact on, or be impacted by, technology R&D investment decisions in each area of sustainability. As previously discussed, many current models restrict the concept of sustainability to include only effects on environmental, social, or economic factors. However, firms operating both in the domestic or global marketplace need to also be concerned with generating sustainable stakeholder value and, consequently, many other factors need to be considered when developing their R&D investment portfolios. These factors include, but are not limited to, the current and projected technology R&D portfolio, expected R&D results, improved operations management processes, and domestic and global business success, in addition to environmental, social, and economic value creation. Whereas governmental regulations and potential business practices may encourage and extract value from "green management" philosophies, these same initiatives may also impose restrictions that further constrain businesses. It is proposed here that broader domains of business, environmental, social, and economic sustainability are essential for the long-term health of domestic and global industrial economics and societal welfare; this comprehensive sustainability must be considered in any strategic technology R&D investment model from each of these perspectives. Therefore, it is necessary to identify and link together each factor of sustainability and develop a conceptual, endogenously-driven model that presents both challenges and business opportunities that are within this expanded view of sustainability.

For example, consider the difficult decisions firms in the automotive industry face when developing technology and innovation R&D investment strategies pertaining to traditional gasoline versus hybrid powered vehicles. Some of the factors contributing to expected revenues, costs, and underlying technological uncertainties include: profits; consumer payback period; R&D costs; governmental regulations; environmental impacts and tradeoffs; and battery production and disposal. From an operations perspective, consider the subsequent effects on global supply chains, including: manufacturing practices; transportation modes; profits; material selection and transportation; forward and reverse logistics; energy usage; and waste generation and disposal. These examples present a sampling of the complexities firms face when trying to achieve overall sustainability while making strategic R&D investment decisions and the necessary improvements to the operations management processes.

This proposed decision-making framework involves the identification of all pertinent parameters and interlinking correlation effects among the following categories of sustainability: business (e.g., technology, innovation, and operations management), environmental, social, and economic sustainability.

4.1 Business Sustainability

Any decision that an organization makes must be geared at achieving both short- and long-term sustainable stakeholder value creation. This value creation may be in terms of profits or setting

the organization up to remain flexible and adjust to future market dynamics. Effective technology R&D investment strategies, the creation of a robust technology portfolio, and the development of appropriate operations management processes are key components for achieving long-term business value creation. For the focus of this model, business sustainability represents the primary good or service that an organization produces, and the required processes in the respective value chain. Therefore, business sustainability may be further segmented into the areas of: (1) technology and innovation management, and (2) operations management.

4.1.1 Technology and Innovation Management

As proposed in [19], organizations should consider the following three technology and innovation management domains: independent, global, and unforeseen industries. The independent industries are the ones that allow a country to promote broad-based sustainability and reduce its dependence on foreign nations. Examples of these independent industries may include the production of food, clothing, and shelter, as well as basic services such as health care. The global industries are those that a nation may invest in to develop key sustainable core competencies that promote its success in the global marketplace. These global industries may either involve all of the R&D efforts to occur within the boundaries of a single country, or perhaps employ a globally-linked network where multiple countries collaborate and provide their own unique core competencies [20]. Possible examples of global industries include energy development, biotechnology, telecommunications, or transportation industries. Finally, the unforeseen industries are the ones that emerge over time and are difficult to predict far in advance. For example, the emergence of the Internet would have been difficult to predict in the 1980s, but successful organizations had to be flexible enough to incorporate this technology into their business strategies and operations processes. Strategic planning with respect to unforeseen industries is especially difficult to predict and model in advance, and also has a major effect on a firms' ability to remain sustainable from a business perspective.

Additional factors that affect business sustainability in the area of technology and innovation management include the following: new product and process developments; discontinuous technologies; establishing a dominant design; cannibalism of current technologies; suffering from incumbent inertia and core rigidity; actions by competitive organizations; development of incremental and/or radical technologies; effects of timing of entry strategies; ability to enter new markets; and the production and availability of enabling technologies or complementary goods [21]. Also, new technologies may be developed that assist with efforts in the areas of operations management, environmental, social, and economic sustainability. As a result, these endogenous effects should be included in any technology R&D investment decision-making model.

4.1.2 Operations Management

As mentioned in [19], operations management processes and infrastructure may be used to assist with new technology and innovation R&D efforts, or they may be simply part of the transformation process in the firms' production of goods and/or services throughout the value chain. As a result, it is important for any organization to continually improve its operations processes and infrastructure to promote sustainable value creation. Continued efforts in the areas of waste reduction and "cradle-to-grave" analyses of supply chains and reverse logistics may help an organization to achieve both economic and environmental sustainability goals. These efforts serve to reduce a firm's pollution costs and metrics, as well as include such concepts as "design for manufacturing, recycling, and environment" into product and process design. As previously mentioned, the development and adherence to the ISO 9000, 14000, and 26000 series provide operational goals that an organization may strive to achieve.

For example, the rapidly-changing electronics industry has created numerous challenges to firms when trying to predict the expected new product development costs. These costs may include design, manufacturing, and forward and reverse logistics as the firm may assume responsibility for the production and delivery phases, as well as the disposal costs (i.e., costs incurred throughout the entire product life cycle). A well-designed operation that adheres to these principles and implements new process innovations should be able to continually improve the efficiency of its transformation process and promote business sustainability for the stakeholders in the area of operations management.

4.2 Environmental Sustainability

In addition to making technology R&D investment decisions that promote sustainable business value creation (i.e., in the proposed domains of technology, innovation, and operations management), it is also important that these decisions account for the ultimate effects on the more traditional area of environmental sustainability. Models should account for both the costs and potential opportunities related to any pertinent ecological factors including the following: air, water, and/or soil quality; toxic waste; energy usage; greenhouse gas emissions; deforestation; acidification of fisheries; and societal health impacts. Depending on the scope of the decision, each of these factors may be represented as either a negative cost constraint or potential source of savings or revenue. For example, a change in the choice of operating energy source to one with cleaner emissions may have a positive effect on environmental sustainability and reduce pollution costs; however, if the additional energy costs exceed the savings, then there will be an ultimate negative effect on business performance and/or sustainability. The complexities of these potential tradeoffs must be incorporated into the technology R&D investment decision-making model.

4.3 Social Sustainability

Organizational investments in technology R&D projects may have profound effects on the ability for societies to remain socially sustainable. A primary motivation for considering the importance of investing in independent, global, and unforeseen industries is to ensure that nations can create an environment where citizens have all of the basics of food, clothing, and shelter, and also access to such necessities as health care and employment opportunities. Advancements in both primary and supportive technologies affect each of these areas from both a firm and societal level. When balancing the objectives of these stakeholders, an example may include investments in the food industry. The addition of pesticides and fertilizers may increase the yield of crops; however, potential soil demineralization or the human ingestion of toxic chemicals may inhibit this as a long-term practice when considering social sustainability. Therefore, the decision-making model should consider these factors, and also note that achieving business sustainability should help satisfy the goals of social sustainability.

4.4 Economic Sustainability

In a similar manner to social sustainability, ensuring that business sustainability is achieved in the areas of technology, innovation, and operations management should help firms and nations to be economically sustainable. On a macroeconomic scale, sustainable business success may provide additional employment opportunities in the independent, global, and unforeseen industries. On a microeconomic scale, decisions that provide continued organizational profits and value creation will allow firms to remain competitive; microeconomic sustainability is a part of the overall comprehensive business sustainability as defined in this paper. Furthermore, achieving economic sustainability also correlates with the probability of providing social sustainability. Therefore, each of these categories of sustainability should be included in the decision-making model, in addition to their endogenous interlinking effects.

5 DECISION-MAKING FOR COMPREHENSIVE SUSTAINABILITY

It is important for an organization to fully model the effects of any technology R&D investment decision in order to properly evaluate whether or not long-term sustainable value can be created across each area of sustainability. This section provides a description of this overall decision-making process; then, the work from [22] is extended through the theoretical model framework presented in Section 6. For each of the areas of business, environmental, social, and economic sustainability proposed in Section 4, it must be determined whether or not there is expected to be a positive effect on sustainability, a negative effect on sustainability, or no effect on sustainability, in addition to any interlinking and correlation parameter effects (e.g.,

governmental regulations that promote environmental sustainability may restrict operational measures or profit potential, and vice-versa). For any given technology R&D investment opportunity, each of the pertinent sustainability factors and parameters must be identified, including:

- Portfolio of all current technology investments for the organization;
- Costs of the proposed project;
- Costs of the required resources;
- Projected environmental impact and costs related to the project;
- Expected future revenue inflows or benefits;
- Expected potential for future technologies and business value creation; and
- Interlinking and parameter endogenous trade-off relationships (e.g., cannibalization effects; development of tacit knowledge).

Once the pertinent parameters, their expected effects on overall sustainability, and the interlinking and correlation relationships are identified, a proper analytical technique may be developed to help assist with the decision-making process. This analytical technique may include portfolio theory to help one determine an optimal allocation of resources across multiple R&D projects [23] [24] [25], multiscale decision theory to provide linkages between agents (i.e., decision-makers) across various government, organization, and societal hierarchies [26], and options-based decision policies to incorporate flexible valuation and allocation techniques that adjust over time as more information is known about the technology and R&D efforts [27] [28]. The results of this model should be able to provide the decision-maker with the following information:

- Given a certain set of conditions and expected sustainability impacts, should the technology R&D investment decision be made or not? Depending on the complexity of the problem domain, various solution techniques may be employed including the use of one or more operations research tools (e.g., decision analysis; simulation) or even staged investment analyses using real options [27] [28].
- How much capital needs to be invested in order to achieve a positive net sustainability value across each of the defined dimensions? The success of any strategic technology R&D investment decision may be dependent on the level of capital invested in the venture. This type of analysis is similar to the traditional net present value (NPV) breakeven or internal rate of return analyses.
- To assist with the business analysis, the ultimate result of this model should be a financial metric of the expected NPV of the investment that is adjusted to reflect the cumulative positive or negative effects of each area of sustainability (i.e., business, environmental, social, and economic sustainability). If appropriate, this metric may then be converted to an index that measures the organization's level of corporate sustainability. For the purposes of this initial theoretical model framework, the ultimate result will be in the

form of a proposed Total Sustainability Index (*TSI*) value that measures the present value of the proposed investment across each area of sustainability in monetary terms.

Once this strategic decision-making model has been tailored to a specific organization, a more comprehensive quantifiable sustainability factor/index can be developed that may provide an expected measured impact of any business decision on that particular organization's unique effect on its future comprehensive sustainability potential. If a benchmark sustainability index value is established for either the specific organization and/or industry, then decisions can be made that are expected to increase, maintain, or reduce the organization's sustainability index. As proposed in this paper, this index should consider the more traditional areas of environmental, social, and economic sustainability, as well as this more comprehensive business metric of sustainability.

For example, whereas investments in a given new technology R&D project may be expected to yield positive financial NPV measures, the effect of the environmental waste aspects may prove that the technology will have negative effects on the firm's *TSI*. Consequently, altering the raw materials for this particular technology R&D project may yield a positive net effect on its comprehensive *TSI* value. Because a firm is ultimately evaluated based on its shareholder value, this proposed model assists organizations when evaluating potential technology R&D investment decisions based on a comprehensive metric that includes the endogenous relationships of business, environmental, social, and economic sustainability factors.

6 THEORETICAL MODEL FRAMEWORK FOR COMPREHENSIVE SUSTAINABLE VALUE

In an effort to create a comprehensive model that assists with making strategic technology R&D investment decisions while accounting for multiple categories of sustainability, the theoretical framework presented in this section may be employed. This framework supports the decision-making process described in Section 5 and also in [22], but is presented in a general form to allow for modeling flexibility when applied to a specific example or decision scenario. The steps of this theoretical model framework are as follows:

- Identify all sustainability categories c, where $c \in C$ and C is the set of all categories. For example, using the categories presented in Section 4, the set C may contain the following sustainability categories: (1) Technology and Innovation Management; (2) Operations Management; (3) Environmental; (4) Social; and (5) Economic.
- Identify all parameter factors k within each sustainability category c. So, let k represent each parameter factor, where $k \in K_c \; \forall \; c \in C$. Here, K_c is the set of all parameter factors within category c. Each of these parameter factors may have a positive or negative effect

on the organization's sustainability rating and will be affected by the ultimate strategic decision.

- For each parameter factor *k*, where $k \in K_c \; \forall \; c \in C$, identify any interlinking relationships between these factors. So, define $\gamma_{i,j,c}$ to be the interlinking effect of parameter factor *i* on parameter factor *j* within category *c*, and $i,j \in K$. This interlinking function accounts for any endogenous relationships between the factors and may model either positive, negative, or no correlations. As a result, the evaluation of an investment decision may include both cumulative as well as trade-off effects on the pertinent parameter factors.
- For each category *c*, where $c \in C$, identify the interlinking relationships between these categories. So, let $\varphi_{m,n}$ represent the interlinking effect of category *m* on category *n* where $m,n \in C$. Similar to the interlinking effects for the individual parameter factors, $\varphi_{m,n}$ represents the collective positive, negative, or neutral endogenous effects of the parameter factors within each category *m* on the overall sustainability value of category *n*.

For any given technology R&D investment decision, it is necessary to determine the overall respective impact on the organization's *TSI* value. This proposed *TSI* value is inclusive of each sustainability category *c* and the cumulative impacts of each parameter factor *k*, where $k \in K_c \; \forall \; c \in C$. Before the *TSI* value can be determined, both the expected values of each parameter factor *k* and category *c* must be calculated for any time period *t*. First, determine the expected value of each factor *j* within category *c* at time period *t*:

$$E\left[x_{j,c}(t)\right] = f\left\{E\left[x_{i,c}(t-1)\right],\, \gamma_{i,j,c}\right\} \cdot E\left[x_{j,c}(t-1)\right] \;\; \forall \; i,j \in K_c \text{ and } c \in C \qquad (1)$$

Where:

$E\left[x_{j,c}(t)\right]$ = the expected value of parameter factor *j* within category *c* at time period *t*;

$f\left\{E\left[x_{i,c}(t-1)\right],\, \gamma_{i,j,c}\right\}$ = the expected endogenous interlinking effect function of each parameter factor *i* on parameter factor *j* within category *c* at the previous time period ($t - 1$); and

$E\left[x_{j,c}(t-1)\right]$ = the expected value of parameter factor *j* within category *c* at the previous time period ($t - 1$).

For notational purposes, then set $E\left[x_{k,c}(t)\right] = E\left[x_{j,c}(t)\right]$ where $k = j \; \forall \; k, j \in K_c$ and $c \in C$.

Now that the individual factor values have been determined at time period *t* within each category *c*, the expected value of each category *n* at time period *t* may be calculated using the following equation:

$$E[C_n(t)] = f\left\{x_{k,c}(t)\right\} \text{ and } f\left\{E[C_m(t-1)],\, \varphi_{m,n}\right\} \cdot E[C_n(t-1)] \quad \forall \; k \in K_c \text{ and } m,n,c \in C \qquad (2)$$

Where:

$E[C_n(t)]$ = the expected value of category *n* at time period *t*;

$f\{x_{k,c}(t)\}$ = the expected value of factor *k* within category *c* at time period *t*;

$f\{E[C_m(t-1)],\ \varphi_{m,n}\}$ = the expected endogenous interlinking effect function value of each other category *m* on category *n* at the previous time period (*t* – 1); and

$E[C_n(t-1)]$ = the expected value of category *n* at the previous time period (*t* – 1).

For notational purposes, then set $E[C_c(t)] = E[C_n(t)]$ where $c = n\ \forall\ c,n \in C$.

Equations (1) and (2) provide a valuation framework that incorporates multiple parameter factors within multiple categories of sustainability, respectively, while allowing the decision-maker to fully account for the endogenous effects of both individual parameter factors and the collective interlinking effects of each category. Once these impacts have been determined for the specified time horizon (t_0,*T*), the expected value of the *TSI* for the R&D investment decision at the predefined ending time period *T* may be calculated using Equation (3):

$$E[TSI(T)] = \sum E[C_c(T)]\ \forall\ c\ \in C \tag{3}$$

Where:

$E[TSI(T)]$ = the expected Total Sustainability Index value at time *T*; and

$\sum E[C_c(T)]$ = the summation of the expected values of each category *c* at the final time period *T*.

Finally, for the specified time horizon (t_0,*T*), one may use this model to determine the current net present value of the *TSI* for the investment decision at time t_0:

$$TSI(t_0) = \frac{\sum E[C_c(T)]}{(1+r)^T}\ \forall\ c\ \in C \tag{4}$$

Where:

$TSI(t_0)$ = the value of the Total Sustainability Index at time t_0;

$\sum E[C_c(T)]$ = the summation of the expected values of each category *c* at the final time period *T*; and

r = the discount rate.

Once the present value of the *TSI* for the potential technology R&D investment decision has been evaluated using Equation (4), this value may then be compared to a desired sustainability impact value to determine whether or not the project investment decision should be made. For example, if the selected decision criteria is to make any technology R&D investment decision that will yield a positive impact on the organization's comprehensive sustainability index across the modeled categories, then the investment will be made provided that $TSI(t_0) > 0$ (i.e., the

decision will be made to proceed with the investment because there is expected to be a positive net effect on the organization's *TSI* value).

7 CONCLUSIONS

The current domestic and global economic climate has led to an increased awareness of the importance of developing and implementing technological R&D investment strategies that help organizations achieve long-term sustainable value creation. This paper presented many definitions and metrics of sustainability that have been previously proposed; however, using them to make technology R&D investment strategic decisions is quite difficult due to the complexity of the affected domains. Furthermore, a broader definition of sustainability has been presented in this paper that encourages decisions promoting business sustainability in the areas of technology, innovation, and operations management in addition to the more traditional areas of environmental, social, and economic sustainability. The overall decision-making model presented in this paper is designed to assist firms in identifying each pertinent parameter factor across multiple categories of sustainability, incorporating any endogenous interlinking effects between both these parameter factors and overall sustainability categories, and determining the overall Total Sustainability Index (*TSI*) value for a given technology R&D investment decision. The model assists organizations in the evaluation of proposed technology R&D decisions and results in a comprehensive total sustainability metric that may be used to promote long-term value creation across multiple dimensions of sustainability: business (e.g., technology, innovation, and operations management), environmental, social, and economic sustainability. Extensions to this approach may be made that help firms compare these individual decision opportunities to organizational and/or industry benchmarks that will allow for sustainable value creation in both the domestic and global marketplaces.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the Business Council of the Leon Hess Business School at Monmouth University for the generous funding of this research project.

REFERENCES

[1] Meadows, D.H., Meadows, D.L., Randers, J., and Behrens III, W.H. (1972) *The Limits to Growth*. Universe Books.

[2] Brundtland, G. (1987) *Our Common Future: The World Commission on Environment and Development*. Oxford: Oxford University Press.

[3] Werbach, A. (2009) *Strategy for Sustainability: A Business Manifesto*. Boston, MA: Harvard Business Press.

[4] Bell, S. and Morse, S. (2010) *Sustainability Indicators: Measuring the Immeasurable?* (2nd ed.). Washington, DC: Earthscan.

[5] Lynam, J.K. and Herdt, R.W. (1989) 'Sense and sustainability: Sustainability as an objective in international agricultural research', *Agricultural Economics*, Vol. 3, pp. 381-398.

[6] Pearce, D. and Turner, R.K. (1990) *Economics of Natural Resources and the Environment*. Harvester Wheatshear, Hemel Hempstead.

[7] Fresco, L.O. and Kroonenberg, S.B. (1992) 'Time and spatial scales in ecological sustainability', *Land Use Policy*, Vol. 9, pp. 155-168.

[8] WCED (World Commission on Environment and Development) (1987) *Our Common Future*. Oxford: Oxford University Press.

[9] IUCN (World Conservation Union), UNEP (United Nations Environment Programme) and WWF (World Wide Fund for Nature) (1991) *Caring for the Earth: A Strategy for Sustainable Living*, IUCN, Gland, Switzerland.

[10] Epstein, M.J. (2008) *Making Sustainability Work: Best Practices in Managing and Measuring Corporate Social, Environmental and Economic Impacts*. Sheffield, UK: Greenleaf Publishing Limited.

[11] Epstein, M.J. and Roy, M-J. (2003) 'Improving sustainability performance: specifying, implementing, and measuring key principles', *Journal of General Management*, Vol. 29, No. 1, pp. 15-31.

[12] Esty, D.C. and Winston, A.S. (2009) *Green to Gold: How Smart Companies Use Environmental Strategy to Innovate, Create Value, and Build Competitive Advantage*. Hoboken, NJ: John Wiley & Sons, Inc.

[13] Ehrlich P. and Holdren, J. (1972) 'A bulletin dialogue on the 'Closing Circle'. Critique: One dimensional ecology', *Bulletin of the Atomic Scientists*, Vol. 28, No. 5, pp. 16–27.

[14] YCELP/CIESIN/ESI (Yale Center for Environmental Law & Policy / Center for International Earth Science Information Network at Columbia University), Available at: http://sedac.ciesin.columbia.edu/es/esi/, Accessed on 2 September 2011.

[15] YCELP/CIESIN/EPI (Yale Center for Environmental Law & Policy / Center for International Earth Science Information Network at Columbia University), Available at: http://sedac.ciesin.columbia.edu/es/epi/, Accessed on 2 September 2011.

[16] Elkington, J. (1998) *Cannibals with Forks*. Gabriola Island, Canada: New Society Publishers.

[17] Atkinson, G. (2000) 'Measuring corporate sustainability', *Journal of Environmental Planning and Management*, Vol. 43, No. 2, pp. 235-252.

[18] Figge, F. and Hahn, T. (2005) 'The cost of sustainability capital and the creation of sustainable value by companies', *Journal of Industrial Ecology*, Vol. 9, No. 4, pp. 47-58.

[19] Ball, D.R. (2010) 'A technology, innovation, and operations strategic model for both domestic and global sustainability', published proceedings from the *2010 Northeast Business and Economics Association (NBEA) Annual Conference*, Morristown, New Jersey, September 30 – October 2, 2010.

[20] Bartlett, C.A. and Ghoshal, S. (1990) 'Managing innovation in the transnational corporation', in C.A. Bartlett, Y. Doz and G. Hedlund (Eds.), *Managing the Global Firm*. London and New York: Routledge.

[21] Schilling, M.A. (2010) *Strategic Management of Technological Innovation* (3rd ed.). New York, NY: McGraw-Hill Irwin.

[22] Ball, D.R. (2011) 'Integrating multiple sustainability criteria in technology, innovation, and operations management strategic decisions', published proceedings from the *2011 Northeast Business and Economics Association (NBEA) Annual Conference*, Philadelphia, Pennsylvania, November 3-5, 2011.

[23] Markowitz, H.M. (1959) *Portfolio Selection: Efficient Diversification of Investment* (Cowles Foundation Monograph 16), New Haven, CT, Yale University Press.

[24] Copeland, T.E. and Weston, J.F. (1992) *Financial Theory and Corporate Policy*, Third Edition, Reading, MA, Addison-Wesley Publishing Company.

[25] Luenberger, D.G. (1997) *Investment Science*, USA, Oxford University Press.

[26] Wernz, C. and Deshmukh, A. (2010) "Multiscale Decision-Making: Bridging Organizational Scales in Systems with Distributed Decision-Makers", *European Journal of Operational Research*, Vol. 202, No. 3, May, pp. 828-840.

[27] Dixit, A.K. and Pindyck, R.S. (1994) *Investment Under Uncertainty*. Princeton, NJ: Princeton University Press.

[28] Trigeorgis, L. (1996) *Real Options: Managerial Flexibility and Strategy in Resource Allocation*. USA: The MIT Press.