Volume 2, Special Issue, Number 1

ISSN 1944-656X H
ISSN 1944-6578 O

BUSINESS STUDIES JOURNAL

Balasundram Manaiam
Sam Houston State University
Special Issue Editor

The official journal of the
Academy for Business Studies,
an Affiliate of the Allied Academies

The *Business Studies Journal* is owned and published by the DreamCatchers Group, LLC. Editorial content is under the control of the Allied Academies, Inc., a non-profit association of scholars, whose purpose is to support and encourage research and the sharing and exchange of ideas and insights throughout the world.

Authors provide the Academy with a publication permission agreement. Neither the Academy, the Allied Academies, nor the DreamCatchers Group is responsible for the content of the individual manuscripts. Any omissions or errors are the sole responsibility of the individual authors. The Editorial Board is responsible for the selection of manuscripts for publication from among those submitted for consideration. The Editors accept final manuscripts in digital form and the Publishers make adjustments solely for the purposes of pagination and organization.

The *Business Studies Journal* is owned and published by the DreamCatchers Group, LLC, PO Box 1708, Arden, NC 28704 USA. Those interested in subscribing to the *Journal*, advertising in the *Journal*, or otherwise communicating with the *Journal*, should contact the Executive Director of the Allied Academies at info@alliedacademies.org.

BUSINESS STUDIES JOURNAL

BUSINESS STUDIES JOURNAL

SPECIAL ISSUE

LETTER FROM THE EDITOR

The *Business Studies Journal* is owned and published by the DreamCatchers Group, LLC. The Editorial Board and the Editors are appointed by the Allied Academies, Inc., a non profit association of scholars whose purpose is to encourage and support the advancement and exchange of knowledge, understanding and teaching throughout the world. The *BSJ* is a principal vehicle for achieving the objectives of the organization.

The *BSJ* is a journal which allows for traditional as well as non-traditional and qualitative issues to be explored. The journal follows the established policy of accepting no more than 25% of the manuscripts submitted for publication. All articles contained in this volume have been double blind refereed.

It is our mission to foster a supportive, mentoring effort on the part of the referees which will result in encouraging and supporting writers. We welcome different viewpoints because in those differences we improve knowledge and understanding.

Information about the Allied Academies, the *BSJ*, and the other journals handled by the Academy, as well as calls for conferences, are published on our web site, www.alliedacademies.org, which is updated regularly. Please visit our site and know that we welcome hearing from you at any time.

Balasundram Maniam, Special Issue Editor
Sam Houston State University

CONCENTRATION AND COMPETITION IN THE BELARUSIAN BANKING INDUSTRY: AN EMPIRICAL ANALYSIS

Vera A. Adamchik, University of Houston-Victoria

ABSTRACT

Economic theory and empirical research provide ambiguous predictions and findings on the effects of concentration on competition. While the structure-conduct-performance paradigm and the efficient structure hypothesis assert a negative trade-off between these two indicators, the growing body of more recent empirical literature (the so called non-structural approach) shows that competitive behavior can exist in very concentrated markets, and collusive behavior can occur in the markets with a large number of banks. So far much of the research on bank concentration and competitiveness has been done for developed countries. Research on this subject matter in the post-communist countries has been scarce, and to our knowledge there is no such research for Belarus, an ex-USSR republic. The paper is one of the first attempts to assess whether a high degree of concentration in the Belarusian banking sector impacted on its competitiveness over 2002-2008. To analyze this issue we calculate a variety of traditional concentration measures as well as a novel measure of competition – the Boone indicator – which assesses the elasticity of a firm's profits with respect to its cost level, with a higher value of this profit elasticity signaling more intense competition. The results show a positive relationship between concentration and competition in the Belarusian banking industry.

INTRODUCTION

It has been proven, both theoretically and empirically, that competition is among the key driving factors of quality, efficiency, and innovation in the banking sector; it also facilitates access of firms and households to banking services and external financing, which ultimately affects economic growth in the country (see, for example, Vives, 2001). It is no surprise, then, that the recent waves of bank mergers in the EU and the US as well as around the world have spurred debates about the impact of bank concentration on competition. Historically, concentration in the banking sector seems to have been more tolerated than that in other industries and even considered beneficial due to a presumed 'concentration-stability' link.[1] For instance, out of 111 countries in the OECD (2008) survey, at the end of 2005 94 countries had three-bank concentration ratios above 50 percent, 62 above 70 percent, and 25 above 90 percent.

Economic theory and empirical research, however, provide ambiguous predictions and findings on the effects of concentration on competition. Almost two decades ago, Shaffer (1992) documented the lack of consensus on this subject matter among the financial economists and concluded that the degree to which banking market structure matters for competition and performance is a hotly debated topic. Many years later, this issue still remains understudied even for developed countries. Bikker and Haaf (2002, p. 53) conclude that "in recent years, however, only a limited number of empirical studies have investigated competition and concentration in European banking markets." Shaffer (2004, p. 288) stresses rapid consolidation among banks in the US and Europe and claims that "the degree of banking competition and its association with market concentration is thus a more relevant issue now than in earlier times." Berger *et al*. (2004, p. 445) echo that "more research is clearly needed on the topic of bank concentration and competition" and continue that "one useful direction for future research is likely to be additional focus on developing nations and their problems of credit availability, economic growth, and financial stability."

The paper's major contributions to the field may be described as follows. First, we analyze a trend in the Belarusian banking sector concentration. Second, so far much of the research on bank concentration and competitiveness has been done for developed countries. Research on this subject matter in the post-communist countries has been scarce, and to our knowledge there is no such research for Belarus, an ex-USSR republic.[2] The paper is one of the first attempts to assess whether a high degree of concentration in the Belarusian banking sector impacted on its competitiveness over 2002-2008. Third, in the empirical literature different concentration measures (like the Herfindahl-Hirschman Index, 3-, 5-firm concentration ratios) and performance measures (such as price-cost margins, or Lerner Index) have been used as a measure of competition. However, it has been shown that those measures have severe drawbacks (Tirole, 1988) and do not necessarily indicate the *competitiveness* of the banking system (Baumol *et al*., 1982). In this paper we apply a novel measure of competition – the Boone indicator – which assesses the elasticity of a firm's profits with respect to its cost level, with a higher value of this profit elasticity signaling more intense competition (Boone 2000, 2008; Boone *et al*., 2005, 2007).

The paper is organized as follows. Section 2 reviews the theoretical background, Section 3 describes the data set, Section 4 presents the methodology and main results, and Section 5 concludes.

THEORETICAL BACKGROUND

During the last several decades, the structure-conduct-performance (SCP) paradigm (Bain, 1951) has been the predominant empirical approach in analyzing banking competition. The 'conduct' aspect of the SCP paradigm posits that a market structure (reflected in concentration measures) is a good indicator of the intensity of competition in this market. More specifically, the argument is that there is a negative relationship between the degree of market concentration and the

degree of competition among banks: concentration encourages collusive behavior among banks and, hence, impedes competition in the sector. The 'performance' aspect of the SCP paradigm links higher concentration in the banking market to less competition, higher prices, and higher banks' profits. Regulatory authorities in many countries still widely use the SCP approach in antitrust assessments: competition is typically measured by concentration ratios; and higher prices in more concentrated, less competitive markets are viewed as socially undesirable.

The SCP view was challenged by the efficient structure (ES) hypothesis which provided an alternative interpretation to the empirical evidence consistent with the SCP paradigm (see, for example, Demsetz, 1973; Peltzman, 1977; Berger, 1995). The ES hypothesis argues that a positive relationship between bank profits and market concentration/structure exists because more efficient (i.e., low cost, high productivity, *etc*.) banks are able to increase profits by reducing prices. Lower prices also help those banks to expand their market shares, thus leading to increased market concentration.

To sum up, both the SCP paradigm and ES hypothesis stem from traditional industrial organization theory and belong to the structural approach direction in the literature on the measurement of bank competition. The SCP paradigm asserts that structure causes performance, while according to the ES view performance causes structure.

The non-structural approach was developed in the context of the new empirical industrial organization literature. This approach posits that concentration/structure alone does not provide a particularly good indicator of competitive behavior and that other factors may affect firms' conduct and performance (Baumol *et al*., 1982; a review in Claessens and Laeven, 2004 and in Northcott, 2004). Barriers to entry, costs of exit, general contestability, risk profiles, branch networks, technology, competition from non-bank financial institutions, the presence of foreign banks, insurance companies and active capital markets all can influence the level of competition in the banking sector. For example, the contestability theory argues that the threat of new entrants alone can induce a bank to behave more competitively. Hence, contrary to the SCP paradigm, the non-structural approach does not *a priori* assume that concentrated markets are not competitive. The non-structural approach asserts that competitive behavior can exist in very concentrated markets, and collusive behavior can occur in the markets with a large number of banks. One of the most important advances of non-structural techniques is that they attempt to directly measure bank competitiveness without knowing the type of market structure. The rationale is that the degree of competition in the banking sector can be determined by the observed price-setting behavior of banks and its deviation from competitive pricing.

Conflicting theoretical predictions along with inconclusive and contradictory empirical evidence[3] highlight the complexity of the linkages between bank concentration and competition. A growing body of research, however, suggests that concentration and competition measure different characteristics of the banking system (Claessens and Laeven, 2004). It is in this setting that this paper analyzes issues of concentration and competition in the Belarusian banking industry. To get

a general idea about the banking market in Belarus, we present various traditional measures of concentration which are considered to be part of the structural approach. We further present and discuss results based upon the estimation of the Boone indicator, a non-structural measure of market competitiveness.

DATA

The data were obtained from the National Bank of the Republic of Belarus. Those are annual data showing bank assets, liabilities, profits and losses as of January 1 of 2002 through 2009. The total number of banks in each year varies from 24 to 31. The sample includes 21 banks that operated throughout the entire period, 10 banks were established and 6 banks were shut down within this period.

METHODOLOGY AND RESULTS

Measuring concentration

To assess the degree of concentration in the Belarusian banking industry, we calculate a variety of traditional concentration measures, such as the *k*-bank Concentration Ratios, the Herfindahl-Hirshman Index, the Comprehensive Industrial Concentration Index, the Hannah-Kay Index, the House Index, the Hall-Tideman Index, and the Theil Entropy Measure [see Bikker and Haaf (2002) for a comprehensive review]. The indices are shown in Table 1.

Table 1. Various concentration measures based on total assets

	1.1.2002	1.1.2003	1.1.2004	1.1.2005
Number of banks, N	24	28	30	31
CR1 (Belarusbank)	0.4299	0.4407	0.4246	0.4137
CR3	0.6406	0.6372	0.6583	0.6688
CR4	0.7167	0.7119	0.7332	0.7504
CR5	0.7923	0.7838	0.8032	0.8209
HHI	0.2257	0.2313	0.2228	0.2190
CCI	0.5081	0.5119	0.5055	0.5042
HKI, alpha = 0.005	23.8296	27.7797	29.7514	30.7328
HKI, alpha = 0.25	17.0340	19.0843	20.0291	20.3258
HKI, alpha = 5	2.8713	2.7840	2.9150	3.0093
HKI, alpha = 10	2.5552	2.4855	2.5903	2.6664

Table 1. Various concentration measures based on total assets				
H, alpha = 0.25	0.3520	0.3533	0.3497	0.3514
H, alpha = 1	0.2304	0.2357	0.2277	0.2244
H, alpha = 2	0.2258	0.2314	0.2229	0.2191
H, alpha = 3	0.2257	0.2313	0.2228	0.2190
HTI	0.1605	0.1547	0.1545	0.1575
Entropy	2.9334	2.9596	2.9854	2.9737
Coef. of variation	2.1018	2.3405	2.3843	2.4062
	1.1.2006	1.1.2007	1.1.2008	1.1.2009
Number of banks, N	30	27	27	31
CR1 (Belarusbank)	0.4316	0.4378	0.4049	0.4044
CR3	0.6841	0.7152	0.6972	0.7162
CR4	0.7649	0.7963	0.7822	0.7839
CR5	0.8378	0.8699	0.8590	0.8512
HHI	0.2353	0.2501	0.2261	0.2361
CCI	0.5237	0.5458	0.5200	0.5355
HKI, alpha = 0.005	29.7338	26.7480	26.7677	30.7001
HKI, alpha = 0.25	19.4061	17.1908	17.6368	19.1840
HKI, alpha = 5	2.8530	2.7955	3.0767	3.0506
HKI, alpha = 10	2.5438	2.5036	2.7308	2.7328
H, alpha = 0.25	0.3721	0.3993	0.3736	0.3896
H, alpha = 1	0.2413	0.2579	0.2333	0.2448
H, alpha = 2	0.2354	0.2504	0.2263	0.2364
H, alpha = 3	0.2353	0.2501	0.2261	0.2361
HTI	0.1690	0.1904	0.1806	0.1802
Entropy	2.8678	2.7145	2.8220	2.7903
Coef. of variation	2.4614	2.3986	2.2595	2.5141

We start our empirical analysis with looking at the *k*-bank concentration ratios. These are the most frequently used measures of concentration due to their simplicity and limited data requirements. The indices take the form:

$$CR_k = \sum_{i=1}^{k} s_i \qquad (1)$$

where k is the number of the largest banks in the market, and s_i is the market share of bank i. The share of the largest bank (Belarusbank) in total assets slightly decreased from about 0.4299 to 0.4044 over 2002-2008. However, the shares of the three, four and five largest banks all increased – from 0.6406 to 0.7162, from 0.7167 to 0.7839, and from 0.7923 to 0.8512, respectively – despite the fact that the total number of banks rose from 24 to 31 from January 2002 to January 2009. The international comparison (OECD, 2008) shows that the Belarusian banking industry was highly concentrated as compared to other transition (post-soviet) and developed countries. For example, among the ex-USSR countries, only Estonia had a higher CR5 ratio of 0.98, while Russia – the closest political and economic ally of Belarus – had the CR5 of 0.438. The other post-communist countries neighboring Belarus had CR5 of 0.8129 (Lithuania), 0.673 (Latvia), 0.486 (Poland). Considering developed countries, Belarus was quite similar to Canada whose CR5 was 0.874 and the total number of banks was 20 domestic and 27 foreign.

In contrast to the k-bank concentration ratios, the Herfindahl-Hirshman Index (HHI) takes into account the entire distribution of bank sizes, incorporates each bank individually, and is calculated as:

$$HHI = \sum_{i=1}^{n} s_i^2 \qquad (2)$$

In the HHI, banks' shares are used as their own weights. Consequently, the HHI assigns a greater weight to larger banks and, hence, stresses their importance in calculating the concentration index. The HHI ranges between $1/n$ (when all banks are of equal size) and 1 (for a monopoly). Over 2002-2008 the HHI exhibited both increasing and decreasing patterns, but overall the HHI increased from 0.2257 to 0.2361. According to the US Department of Justice, the Belarusian banking sector would be classified as concentrated, and an increase in the HHI by 0.0104 would raise antitrust concerns.[4]

Both the k-bank concentration ratios and the HHI suggest that the Belarusian banking industry became more concentrated. To visualize the process, the two concentration curves for the Belarusian banking industry as of January 1, 2002 and January 1, 2009 were drawn (not shown in this paper). Each curve plotted the cumulative market share in total assets against the number of banks. As expected, the figure confirmed an increased degree of concentration in the banking industry.

As noted in Bikker and Haaf (2002, p. 63), "despite the widely recognized convention that the dominance of the largest few banks determines market behavior, discrete concentration measures have been criticized on the grounds that they ignore changes in market structure occurring elsewhere than among the largest banks." Horvarth (1970) presented the Comprehensive Industrial Concentration Index (CCI), which was designed to reflect both absolute magnitude and relative dispersion:

$$CCI = s_1 + \sum_{i=2}^{n} s_i^2 (1 + (1 - s_i)) \quad (3)$$

where s_1 is the share of the biggest bank, and (1 + (1 - s_i)) is the weight for bank *i*, reflecting the share of the rest of the industry. The CCI approaches zero for an infinite number of equally sized banks and unity for a monopoly. Our calculations show that the CCI increased from 0.5081 to 0.5355 over 2002-2008.

Hannah and Kay (1977) proposed to use a deliberately chosen elasticity parameter α (α > 0, α ≠ 1) to define the appropriate weighting scheme which would emphasize either the lower or upper portion of the bank distribution:

$$HKI = \left(\sum_{i=1}^{n} s_i^{\alpha} \right)^{\frac{1}{1-\alpha}} \quad (4)$$

As Table 1 shows, for $\alpha \to 0$ (0.005 in our calculations), the HKI approaches the number of banks in the industry. For $\alpha \to \infty$ (10 in our calculations), the HKI approaches 1/the share of the largest bank. Despite that fact that the number of banks increased from 24 on January 1, 2002 to 31 on January 1, 2009, the two HKI indices (for α = 0.25 and 5) increased from 17.0340 to 19.1840 and from 2.8713 to 3.0506, respectively. Higher concentration implies that the size effect outweighed the number effect.

House (1977) introduced a parameter α reflecting the degree of collusion, with low values of α implying a high degree of collusion:

$$H = \sum_{i=1}^{n} s_i^{2-(s_i(HHI - s_i^2))^{\alpha}} . \quad (5)$$

The index approaches zero for an infinite number of equally sized banks and unity for a monopoly when α = 0.25 (that is, assuming a highly collusive market), the House index for the Belarusian banking sector grew from 0.3520 to 0.3896 from January 2002 to January 2009 with α = 2 and 3 (that is, assuming a non-collusive market), the House index grew from 0.2258 to 0.2364 and from 0.2257 to 0.2361, respectively. In the latter case, the House index converged to the Herfindahl-Hirshman result.

Hall and Tideman (1967) believed that the number of banks should be included in the calculation of the concentration index in order to reflect the conditions of entry into the industry. Their index takes the form:

$$HTI = \frac{1}{2\sum_{i=1}^{n} is_i - 1} \quad (6)$$

where i is the bank's rank, with the largest bank receiving i=1. The HTI approaches zero for a big number of equally sized banks and unity for a monopoly. In our case, the HTI ranges between 0.15 and 0.19. Over the 2002-2008 period the HTI increased from 0.1605 to 0.1802, indicating an increase in concentration.

The next measure of concentration used in this paper is that introduced by Theil (1967). The Entropy measure was adopted from thermodynamics into information theory and then into economics. It measures the expected information content of a distribution:

$$E = -\left(\frac{1}{\ln 2}\right)\sum_{i=1}^{n} s_i \ln s_i \quad (7)$$

Unlike all other concentration measures discussed above, the Entropy index varies inversely with the degree of concentration, and ranges between 0 (for a monopoly) and $log_2 n$ (for equally sized banks). The Entropy indices decreased from 2.9334 to 2.7903 over 2002-2008, implying an increase in concentration.

To assess the dispersion of total assets in the banking sector, we augment our brief assessment of concentration with the coefficient of variation:

$$CV = \frac{\sqrt{Var(x)}}{\bar{x}} \quad (8)$$

where $\bar{x}$ is the mean bank size. If the sizes of all banks increase proportionally, the coefficient of variation will remain unchanged. The calculated coefficient of variation exhibits an increasing trend, with the values of 2.1018 on January 1, 2002 and 2.5141 on January 1, 2009. It indicates that the dispersion of the bank sizes around the mean increased, which together with all other concentration indices in Table 1 suggest that there had been an increased in the concentration of total financial assets under the control of Belarus' largest banks in 2002-2008. Our next step is to investigate whether this development in the market structure led to a less (more) competitive behavior of the banks.

Measuring competition

To measure the level of competition in the banking sector we use the Boone indicator (Boone 2000, 2008; Boone *et al*., 2005, 2007). The Boone indicator assesses the relationship between performance, in terms of profits, and efficiency, measured as marginal costs. Typically, in any market, efficient firms have higher profits than inefficient firms. However, in a more competitive market, efficient firms are rewarded more and inefficient firms punished more harshly (in terms of profits) than they are in uncompetitive markets. Hence, Boone suggests measuring the competitiveness of a market by estimating the elasticity of a firm's profits with respect to its cost level. The expected sign of the β coefficient is negative, and a more negative β indicates more intense competition.[5] Roughly speaking, the following specification is estimated:

$$\ln \pi = \alpha + \beta \ln mc + \varepsilon_{\pi} \tag{9}$$

where π is profits; *mc*is marginal cost; and the slope β is interpreted as the profit elasticity. Since it is impossible to directly observe marginal costs, some researchers approximate marginal costs using average variable costs (Boone *et al*., 2005, 2007) and some researchers calculate marginal costs using a cost function (Leuvensteijn *et al*., 2007). In this paper, we follow the latter approach and first estimate a translog cost function for the Belarusian banking sector using individual bank observations. Due to its flexibility of specification, a translog cost function has been extensively employed in many studies of depository institutions.

In specifying the cost function, we rely on the intermediation model of a bank, as developed by Klein (1971) and Sealey and Lindley (1977).[6] This approach views the bank as a firm collecting deposits and other funds in order to transform them into loans and other assets. For this transformation, physical capital and labor are employed. Hence, the major inputs in the bank production process are deposits and other funds, labor, and physical capital; and the output is typically measured by loans and other income generating activities (banking services). As actual factor price data are not available, we proxy them by ratios of expenses to respective volume following the literature. We estimate a translog cost function with one output (loans)[7], three inputs (funds, labor, and physical capital), one control variable, and annual dummies:

$$\begin{aligned}\ln C = {} & \alpha + \beta_1 \ln p_1 + \beta_2 \ln p_2 + \beta_3 \ln p_3 + \beta_y \ln Y \\ & + \delta_{11} \frac{1}{2} \ln^2 p_1 + \delta_{22} \frac{1}{2} \ln^2 p_2 + \delta_{33} \frac{1}{2} \ln^2 p_3 + \beta_{yy} \frac{1}{2} \ln^2 Y \\ & + \delta_{12} \ln p_1 \ln p_2 + \delta_{13} \ln p_1 \ln p_3 + \delta_{23} \ln p_2 \ln p_3 \\ & + \gamma_{y1} \ln Y \ln p_1 + \gamma_{y2} \ln Y \ln p_2 + \gamma_{y3} \ln Y \ln p_3\end{aligned}$$

$$+\theta_e \ln\frac{E}{A}+\theta_{ee}\frac{1}{2}\ln^2\frac{E}{A}+\sum_{t=1}^{T-1}\phi_t d_t+\varepsilon_c \quad (10)$$

where

C is total bank expenses;
Y is loans to clients and other banks;
p_1 is price of labor, proxied by administrative expenses (the predominant portion of which is personnel expenses) to total assets;
p_2 is price of funding, proxied by interest expenses divided by total funds;
p_3 is price of fixed capital, proxied by depreciation expenses divided by fixed assets;
$\frac{E}{A}$ is the equity to assets ratio used as a control variable to correct for differences in loan portfolio risk across banks (see Berger and Mester, 1997);
d_t are the binary time dummy variables which are designed to capture technological change. They also intend to absorb the impact of inflation on our results because in our analysis we use nominal values[8].

The cost shares of funds, labor, and capital are given by:

$$S_1=\beta_1+\delta_{11}\ln p_1+\delta_{12}\ln p_2+\delta_{13}\ln p_3+\gamma_{y1}\ln Y+\varepsilon_1 \quad (11)$$

$$S_2=\beta_2+\delta_{12}\ln p_1+\delta_{22}\ln p_2+\delta_{23}\ln p_3+\gamma_{y2}\ln Y+\varepsilon_2 \quad (12)$$

$$S_3=\beta_3+\delta_{13}\ln p_1+\delta_{23}\ln p_2+\delta_{33}\ln p_3+\gamma_{y3}\ln Y+\varepsilon_3 \quad (13)$$

The full model includes Eqns. (10)-(13). By construction,

$$\gamma_{yi}=\gamma_{iy},\ \delta_{ij}=\delta_{ji},\ \sum_{i=1}^{3}S_i=1,\ \sum_{i=1}^{3}\beta_i=1,\ \sum_{i=1}^{3}\delta_{ij}=0,\text{ and }\sum_{j=1}^{3}\delta_{ij}=0 \quad (14)$$

These conditions can be imposed directly on the model by specifying the translog model in (C/p_3), (p_1/p_3), and (p_2/p_3) and dropping the third share equation (13). Now the full model consists of Eqns. (15)-(17):

$$\ln\frac{C}{p_3}=\alpha+\beta_1\ln\frac{p_1}{p_3}+\beta_2\ln\frac{p_2}{p_3}+\beta_y\ln Y$$
$$+\delta_{11}\frac{1}{2}\ln^2\frac{p_1}{p_3}+\delta_{22}\frac{1}{2}\ln^2\frac{p_2}{p_3}+\beta_{yy}\frac{1}{2}\ln^2 Y$$
$$+\delta_{12}\ln\frac{p_1}{p_3}\ln\frac{p_2}{p_3}+\gamma_{y1}\ln Y\ln\frac{p_1}{p_3}+\gamma_{y2}\ln Y\ln\frac{p_2}{p_3}$$
$$+\theta_e\ln\frac{E}{A}+\theta_{ee}\frac{1}{2}\ln^2\frac{E}{A}+\sum_{t=1}^{T-1}\phi_t d_t+\varepsilon_{c/p_3} \quad (15)$$

$$S_1=\beta_1+\delta_{11}\ln\frac{p_1}{p_3}+\delta_{12}\ln\frac{p_2}{p_3}+\gamma_{y1}\ln Y+\varepsilon_1 \quad (16)$$

$$S_2=\beta_2+\delta_{12}\ln\frac{p_1}{p_3}+\delta_{22}\ln\frac{p_2}{p_3}+\gamma_{y2}\ln Y+\varepsilon_2 \quad (17)$$

The model above reduces the number of estimated parameters from 23 to 18. The rest of parameters is estimated using Eqns. (14). We estimate the full model (15)-(17) by maximum likelihood to ensure invariance with respect to the choice of which share we drop.

The cost function in Eqn. (10) implies a marginal cost function of the form:

$$mc=\left(\frac{C}{Y}\right)*\frac{\partial\ln C}{\partial\ln Y}=\left(\frac{C}{Y}\right)*(\beta_y+\beta_{yy}\ln Y+\gamma_{y1}\ln p_1+\gamma_{y2}\ln p_2+\gamma_{y3}\ln p_3) \quad (18)$$

We estimate marginal costs for all bank observations and then regress them on total gross (before-tax) profits of each bank as shown in Eqn. (9). The relative profits measure (i.e., Boone indicator) is captured by the estimated coefficient β.

In our estimations, we use average annual values of assets, liabilities and their categories, calculated as a simple mean of their values reported on January 1 of the year under consideration and January 1 of the next year. This reduces the total number of banks in each year to 21-27. The sample includes 21 banks that operated throughout the entire period and 6 banks that were established within this period. The total number of bank-year observations is 180.

The estimates of the translog cost function are shown in Table 2. Marginal costs at the individual bank level were calculated using Eqn. (18). The dynamics of average marginal costs of loans during 2002-2008 is shown in Table 3. For each year, individual marginal costs were weighted by the amount of loans on a bank's balance sheet. Table 3 clearly shows that average marginal costs in the Belarusian banking sector gradually declined from 29 percent to 13 percent over 2002-2008, which mainly reflects the decrease in funding rates over this period.

Table 2. Estimates of the translog cost function							
Coefficient	Value	Standard Error	b/St.Er.	Coefficient	Value	Standard Error	b/St.Er.
α	-1.1171	0.3472	-3.218	β_{YY}	-0.0207	0.0050	-4.150
β_1	0.8739	0.0266	32.896	θ_e	-0.6149	0.0822	-7.483
β_2	0.0947	0.0289	3.281	θ_{ee}	-0.2101	0.0248	-8.464
β_Y	1.2593	0.0613	20.555	φ2003	0.0047	0.0469	0.100
δ_{11}	0.1465	0.0060	24.450	φ2004	0.0244	0.0463	0.527
δ_{12}	-0.1494	0.0062	-24.097	φ2005	0.0714	0.0459	1.554
δ_{22}	0.1559	0.0068	22.898	φ2006	0.0678	0.0462	1.470
γ_{y1}	-0.0242	0.0023	-10.753	φ2007	0.0276	0.0475	0.580
γ_{Y2}	0.0243	0.0024	9.937	φ2008	0.0097	0.0486	0.199

We next estimated the Boone indicator for the entire 2002-2008 period and for each year separately. The results are presented in Table 3. For the full sample period, the Boone indicator is negative (as expected), statistically significant but rather small in economic terms. The estimated β of -0.74 suggests that a bank with 1 percent higher marginal costs than another (more efficient) bank would have 0.74 percent lower profits than the more efficient bank. For comparison, we refer to the studies by Leuvensteijn *et al*. (2007) and Maslovych (2009) who also estimated the Boone indicators from a translog cost function. Leuvensteijn *et al*. (2007) estimate the Boone indicator for 8 developed countries over 1994-2004 and report the highest value of -5.41 for the U.S., followed by -4.15 for Spain, -3.71 for Italy, -3.38 for Germany, -1.56 for the Netherlands, -1.05 for the UK, -0.90 for France, and -0.72 for Japan. This international comparison suggests that Belarusian banks are less competitive, as compared to the U.S. and the euro area. The degree of competition in the Belarusian banking sector seems to be similar to that in Japan. Maslovych (2009) reports the Boone indicators for Ukraine, a post-Soviet transition country, for 2006-2008. For the entire period, the estimated Boone indicator is -1.61, implying that the Ukrainian banking sector is more competitive that the Belarusian one.

Table 3. Marginal costs of loans and the Boone indicator over time					
Period	Marginal costs of loans, % of loans	The Boone indicator			
			Standard Error	*t*-ratio	P[|T|>*t*
2002	28.9	1.1262	0.7294	1.544	0.1391
2003	22.4	-0.0309	0.6998	-0.044	0.9652
2004	19.0	0.3366	0.7748	0.434	0.6678
2005	15.4	-1.7405	1.0949	-1.590	0.1250
2006	13.7	-1.6319	0.8957	-1.822	0.0810
2007	13.1	-2.4206	0.6788	-3.566	0.0015
2008	13.0	-2.3758	0.5514	-4.308	0.0002
2002-2008		-0.7435	0.3026	-2.457	0.0150

Overall, the Boone indicators calculated for the entire sample period may conceal considerable differences over time. The annual Boone indicators in Table 3 show a decreasing trend (that is, indicating an increase in competition). For 2002-2004, the Boone values are not statistically significant; moreover for 2002 and 2004 the values are positive, which is against our expectations. However, starting 2005, the Boon indicators are negative, statistically significant, and exhibit a decreasing trend with relatively high values of competition for the most recent years – about -2.4 for 2007 and 2008. The plots for 2007 and 2008 (not shown in this paper) demonstrate that log profits are decreasing in log marginal costs; or, in other words, banks with higher marginal costs earn lower profits. International comparison suggests that Belarus, with its value of -1.7 in 2004, fits quite well into the distribution of the 8 developed countries in Leuvensteijn *et al*. (2007). For 2004 (the most recent year in their study), Leuvensteijn *et al*. report the following values of the Boone indicator: -4.54 for the U.S., -3.63 for Japan, -3.09 for the Netherlands, -2.69 for Spain, -2.66 for Germany, -1.81 for Italy, -0.49 for the UK, and 0.10 for France. For Ukraine, Maslovych (2009) reports -1.24 for 2006, -1.15 for 2007, and -2.29 for 2008, which is comparable with our findings for Belarus.

Finally, we calculated correlation coefficients for the Boone indicator and different concentration measures shown in Table 1. For convenience, we used the negative values of the Boone indicator and of the Entropy measure, so that a positive correlation coefficient indicates an increase in both concentration and competition. With only a few exceptions, the results shown in Table 4 are positive implying that in 2002-2008 both concentration and competition in the Belarusian banking market increased. The positive correlation may be caused by a complex relationship between bank concentration and the measure of bank competitiveness calculated from marginal bank behavior. It may be the case that various factors related to market structure (such as

institutional framework, regulation, contestability, foreign entry, and macro-economic stability) differently affect the competitive climate in the banking sector. It is also plausible that small-sized banks were able to provide meaningful competition to the five largest banks. To conclude, our finding contradicts the SCP paradigm and suggests that one should not focus solely on concentration indices as a measure for competition and that a concentrated market may be competitive.

Table 4. Correlation coefficients for the Boone indicator and different concentration measures

Concentration measure	Correlation coefficient	Concentration measure	Correlation coefficient
A negative value of the Boone indicator	1.00000	A negative value of Entropy	0.69000
CR1 (Belarusbank)	-0.34965	H, alpha = 0.25	0.73033
CR3	0.89097	H, alpha = 1	0.47676
CR4	0.91057	H, alpha = 2	0.41064
CR5	0.89965	H, alpha = 3	0.40805
HHI	0.40805	HKI, alpha = 0.005	0.31361
CCI	0.63409	HKI, alpha = 0.25	-0.04894
HTI	0.74254	HKI, alpha = 5	0.32360
Coef. of variation	0.52837	HKI, alpha = 10	0.35977

CONCLUSIONS

This paper has attempted to assess whether high concentration in the Belarusian banking sector impacted on its competitiveness over 2002-2008. To analyze this issue we calculated various traditional concentration measures and the novel measure of competition (the Boone indicator). The results show a positive relationship between concentration and competition.

ENDNOTES

[1] Proponents of this view rely upon the following arguments. First, they argue that competition leads to a decline in bank efficiency (primarily profit efficiency). The rationale behind this hypothesis is that competitive environment increases customers' propensity to switch to other banks, and the bank-customer relationships become shorter and unstable. Consequently, banks have to spend additional resources for screening, monitoring, attracting, and retaining their clients. Also, banks are likely to experience a greater share of non-performing loans and incur losses. Lower profits in more competitive markets make the banking system more fragile and vulnerable to adverse shocks, while higher profits in less competitive markets provide a 'capital buffer' against such shocks, increase a bank's 'franchise value' and deter risk-taking behavior of the bank's management. Second, some economists argue that banks in a less competitive, more concentrated market tend to be larger,

more diversified, and hence less risky. Third, since it is substantially easier to monitor only a few banks, supervision of banks will be more effective and the risks of a bank failure less pronounced in a concentrated banking industry.

[2] Overall, for whatever reason, Belarus is rarely included into the analyses of development and performance of the banking sector in transition countries. We are aware of the following small set of studies analyzing the Belarusian banking industry either separately or in a cross-country setting: Fries *et al*. (2002), Daneyko and Kruk (2005), Minuk *et al*. (2007), Delis (2009), Delis and Pagoulatos (2009).

[3] For instance, Fernandez de Guevara *et al*. (2005) do not find any significant relation between concentration and competition; Bikker and Haaf (2002) find a negative relationship, and Claessens and Laeven (2004) find a positive relationship.

[4] According to the US Department of Justice, markets in which the HHI is between 1000 and 1800 points (0.1-0.18) are considered to be moderately concentrated, and those in which the HHI is in excess of 1800 points (0.18) are considered to be concentrated. Transactions that increase the HHI by more than 100 points (0.01) in concentrated markets presumptively raise antitrust concerns under the Horizontal Merger Guidelines issued by the US Department of Justice and the Federal Trade Commission. (http://www.justice.gov/atr/public/testimony/hhi.htm).

[5] It is worth to note that "it is not necessarily the case that an increase in competition reduces *every* firm's profits. (…) an increase in competition increases profits of a firm *relative* to a less efficient firm. (…) The benchmark firm could be the median firm or the least efficient firm in the market. The exact identity of this firm does not matter as it will end up in the time fixed effects." For more explanations and derivations see Boone *et al.*, 2007, p. 43.

[6] There are two major ways how the production process in banking is described in the literature: the 'production approach' and the 'intermediation approach.' For empirical purposes, the crucial difference between these two approaches lies in their treatment of deposits. The intermediation approach considers deposits as an input factor, while the production approach considers deposits as an output. See, for example, Mlima and Hjalmarsson (2002) for an overview and comparison of different studies.

[7] We could also extend our model to multiple products in order to estimate a separate degree of competition for each product segment. However, in our sample, loans is the only output category produced by all banks. Many banks do not work with securities or investments.

[8] In some studies, variables are deflated by the GDP Deflator. Shaffer (1990), however, found no qualitative difference between real, nominal, and hybrid specifications.

REFERENCES

Bain, J. (1951). Relation of profit rate to industry concentration. *Quarterly Journal of Economics*, 65, 293-324.

Baumol, W., J. Panzar & R. Willig (1982). *Contestable markets and the theory of industry structure.* San Diego: Harcourt Brace Jovanovich.

Beck, T., A. Demirguc-Kunt & R. Levine (2006). Bank concentration, competition, and crises: First Results. *Journal of Banking and Finance*, 30, 1581-1603.

Berger, A. (1995). The profit-structure relationship in banking: Tests of market-power and efficient-structure hypotheses. *Journal of Money, Credit, and Banking*, 27, 404-431.

Berger, A. & L. Mester (1997). Inside the black box: What explains differences in the efficiencies of financial institutions? *Journal of Banking and Finance*, 21, 895-947.

Berger, A., A. Demirguc-Kunt, R. Levine & J. Haubrich (2004). Bank concentration and competition: An evolution in the making. *Journal of Money, Credit and Banking*, 36, 433-451.

Bikker, J. & K. Haaf (2002). Competition, concentration and their relationship: An empirical analysis of the banking industry. *Journal of Banking and Finance*, 26, 2191- 2214.

Bikker, J. & K. Haaf (2002). Measures of competition and concentration in the banking industry: A review of the literature. *Economic and Financial Modelling*, 9 (summer), 53-98. Reprinted in: Bikker, J. (2004). *Competition and efficiency in a unified European banking market*. Edward Elgar.

Boone, J. (2000). Competition. *CEPR Discussion Paper* No. 2636.

Boone, J. (2008). A new way to measure competition. *Economic Journal*, 118, 1245-1261.

Boone, J., R. Griffith & R.Harrison (2005). Measuring competition. *AIM Research Working Paper* No. 022.

Boone, J., J. Van Ours & H. van der Wiel (2007). How (not) to measure competition. *TILEC Discussion Paper* No. 2007-014; *CentER Discussion Paper* No. 2007-32.

Claessens, S. & L. Laeven (2004). What drives bank competition? Some international evidence. *Journal of Money, Credit and Banking*, 36, 563-584.

Daneyko, P. & D. Kruk (2005). Reforming the banking system of Belarus. *Problems of Economic Transition*, 48, 68-95.

Delis, M. (2009). Competitive conditions in the Central and Eastern European banking systems. *Omega*, in press.

Delis, M. & G. Pagoulatos (2009). Bank competition, institutional strength and financial reforms in Central and Eastern Europe and the EU. *MPRA Paper* No. 16494.

Demsetz, H. (1973). Industry structure, market rivalry, and public policy. *Journal of Law and Economics*, 16, 1-9.

Fernandez de Guevara, J., J. Maudos & F. Perez (2005). Market power in European banking sectors. *Journal of Financial Services Research*, 27, 109-137.

Fries, S., D. Neven & P. Seabright (2002). Bank performance in transition economies. *EBRD Working Paper* No. 76.

Hall, M. & N. Tideman (1967). Measures of concentration. *American Statistical Association Journal*, 62, 162-168.

Hannah, L. & J. Kay (1977). *Concentration in modern industry*. London: MacMillan Press.

Horvarth, J. (1970). Suggestion for a comprehensive measure of concentration. *Southern Economic Journal*, 36, 446-452.

House, J. (1977). The measurement of concentrated industrial structure and the size distribution of firms. *Annals of Economic and Social Measurement*, 6, 73-107.

Klein, M. (1971). A theory of the banking firm. *Journal of Money, Credit, and Banking*, 7, 205-218.

Leuvensteijn, M. van., J. Bikker, A. van Rixtel & C. Sorensen (2007). A new approach to measuring competition in the loan markets of the Euro area. *European Central Bank Working Paper* No. 768.

Maslovych, M. (2009). *The Boone indicator as a measure of competition in banking sector: The case of Ukraine*. Unpublished master's thesis, Kyiv School of Economics.

Minuk, O., F. Rossaro & U. Walther (2007). Deposit insurance reform in Belarus: Remedy or threat for the banking system? *Transition Studies Review*, 14, 22-39.

Mlima, A. & L. Hjalmarsson (2002). Measurement of inputs and outputs in the banking industry. *Tanzanet Journal*, 3, 12-22.

Northcott, C. (2004). Competition in banking: A review of the literature. *Bank of Canada Working Paper* No. 2004-24.

Peltzman, S. (1977). The gains and losses from industrial concentration. *Journal of Law and Economics*, 20, 229-263.

Sealey, C. & J. Lindley (1977). Inputs, outputs and a theory of production and cost at depository financial institutions. *Journal of Finance*, 4, 1251-1266.

Shaffer, S. (1990). A test of competition in Canadian banking. *Federal Reserve Bank of Philadelphia Working Paper* No. 90-18.

Shaffer, S. (1992). Competitiveness in banking. In P. Newman, M. Milgate & J. Eatwell (Eds.), *The New Palgrave Dictionary of Money and Finance* (Vol. 1, pp. 414-416). Stockton Press.

Shaffer, S. (2004). Patterns of competition in banking. *Journal of Economics and Business*, 56, 287-313.

Theil, H. (1967). *Economics and information theory (Studies in mathematical and managerial economics, Vol. 7)*. Amsterdam: North-Holland Publishing Company.

Tirole, J. (1988). *The theory of industrial organization*. MIT Press.

Vives, X. (2001). Competition in the changing world of banking. *Oxford Review of Economic Policy*, 17, 535-547.

World Bank (2008). *Bank regulation and supervision dataset*. Available at http://go.worldbank.org/SNUSW978P0

TEXAS BANKING IN THE ECONOMIC DOWNTURN

James B. Bexley, Sam Houston State University

ABSTRACT

Texas was the last state to realize the economic downturn, and it was generally thought that the financial impact to the economy and the banking system was minimal. The data shows that while the downturn had a major impact on bank profits as well as asset quality in Texas, the impact on the rest of the United States banks was substantially more severe. This study will present a profile of bank indicators and will examine the data from 2005 to the end of 2009 and compare Texas to the United States data.

The areas that are studied are return on assets (ROA), return on equity (ROE) loan charge-offs, allowance for loan and lease losses, non-performing loans to total loans, net interest margin, and unprofitable banks. Additionally, the banks currently under cease and desist orders will be considered. Each of these areas will be explained and then be examined to determine how Texas banks performed in relation to U. S. banks.

INTRODUCTION

The banking industry in the United States is experiencing times that recall the crisis of the late 1980s and early 1990s. Texas banks have had some serious problems, however, they pail by comparison with the rest of the nation. Since January 2008, the Federal Deposit Insurance Corporation failed 181 banks, of those, only 8 were in Texas. Although this is a relatively small percentage of the total banks, the trend that this represents is the real cause for concern. According to SNL Financial Data Dispatch, if you annualize the failures to date since the beginning of 2010 (41 banks), there should be approximately 176 failures in 2010. However, to put the crisis in perspective, with the financial crises of the 1980s and early 1990s, the Federal Deposit Insurance Corporation noted that between 1980 and 1994, they closed 1,600 banks.

CAUSES OF THE BANKING CRISIS

Klomp (2010) studied banks in 110 countries and through coefficient logit modeling found that there was a correlation that high credit growth, a negative growth of GDP, and high interest rates were root causes, but noted that 60 percent of the bank failures were not caused by those three elements rather high economic development was the culprit. Allen and Carletti (2010) argue that the real estate price bubble, loose monetary policy by the Federal Reserve Bank, and global

imbalances caused the banking crisis. In a strange twist, Meadowcroft (2010) dispelled some myths about the crisis. He noted that it was not banker greed rather misguided attempts by bankers to act prudently. This study will look at specific issues in relation to each other to demonstrate the difference in impact on the Texas banking scene and the nation as a whole.

TEXAS AND U. S. BANKING PROFILE

According to the Federal Deposit Insurance Corporation (2010), at the end of 2009, there were 629 banks in the State of Texas with $371,495,000,000 in total assets. Of these banks there were 249 banks under $100 million in asset size, 214 banks with assets of $100 million to $250 million, 125 banks with assets of $250 million to $1 billion, 36 banks with assets from $1 billion to $10 billion, and 5 banks over $10 billion in assets. The F. D. I. C. Quarterly Banking Profile indicated that there were 6,839 banks in the United States at the end of 2009 with assets of $11,109.5 billion. There were 2,525 banks under $100 million in assets, 3,800 banks from $100 million to $1 billion in assets, 429 banks with $1 billion to $10 billion in assets, and 85 banks with assets greater than $10 billion.

BANK INDICATORS

Each of the following indicators impact bank performance and give an indicator of the economic conditions in the state of Texas and the United States. To present the comparison between the state of Texas and the U. S., each category will be examined on an annual basis.

Return on Assets

Over the years, return on assets has been the measure of performance to evaluate bank performance. Good performance has been judged to be at the one percent or above level. During 2005, the Texas banks earned 1.32% ROA and U. S. banks earned 1.30%, both having very good performance. Likewise, in 2006 the Texas banks and U. S. banks performed well, in fact both improved to 1.33%. As the economy started to slow at the end of 2007, performance in Texas banks declined slightly in 2007 reaching the 1.24% level, while the U. S. banks experienced a larger decline to 0.93%. The economy started its major decline in 2008, Texas banks showed some economic slide by dropping to 0.81%, however, the U. S. banks were in free fall declining to 0.13%. Texas improved to 0.86% in 2009 showing some economic improvement in return on assets, but the U. S. banks fell further into near loss at 0.09%.

Return on Equity

The return on equity is the measure of value increase of the stakeholder's investment. A good measure of the equity return is the 10 to 15 percent range. The results of measuring return on equity of Texas banks and U. S. banks are only slightly different from that of the return on assets. In 2005, the Texas banks earned 12.31 ROE and U. S. banks earned 12.87% ROE, both having good performances with the U. S. banks doing better than Texas banks. During 2006 the Texas banks declined slightly to 11.80% and U. S. banks performed better with a 13.02% ROE. As the economy started to slow at the end of 2007, performance in Texas banks declined slightly reaching the 10.23% level, while the U. S. banks experienced a larger decline to 9.12%. The economy started its major decline in 2008, Texas banks showed substantial economic slide by dropping to 5.74%, while, the U. S. banks were in free fall dropping 780 basis points, declining to 1.32%. Texas declined 52 basis points to 5.22% in 2009 and the U. S. banks fell further to 0.85%.

Charge-offs

A charge-off is the portion of a loan that is deemed uncollectable and must be written off the bank's books. Historically, loan charge-offs in good economic times range in the 0.15% to the 0.25% range. For the year of 2005 the Texas banks reflected charge-offs at 0.23% and U. S. Banks stood at 0.56% so in good times Texas had approximately one-half the amount of charge-offs for U. S. banks. Charge-offs declined during 2006 in Texas banks dropping to 0.19% and the U. S. banks declined to 0.41%. Levels of charge-offs continued through 2007 rising only slightly to 0.22% in Texas and in the U. S. to 0.62%. Charge-offs more than doubled in 2008 to 0.49% in Texas and to 1.32% in the U. S. as the economy started its decline. The charge-offs in Texas banks improved substantially in 2009 to 0.21%, but in U. S. banks more than doubled the 2008 level to 2.57%.

Allowance for Loan and Lease Losses

The allowance is a special reserve account set aside to insulate the bank from losses on loans and leases. During normal economic times, an allowance of 1.10% was considered average. In 2005, the allowance for loan and lease losses averaged 1.19% in Texas and 1.12 in the U. S. During 2006, the allowance dropped to 1.11% in Texas and rose to 1.15% in the U. S. Regulators allowed the allowance for loan and lease losses to dip to 1.09% in Texas, believing that the Texas economy was robust, but increased the allowance in U. S. banks to 1.35% seeing some instability nationally. As the economy showed substantial deterioration during 2008, Texas banks were "encouraged" to increase the allowance to 1.32% and U. S. banks to 2.29%. With the economy heading for a

downturn, Texas banks raised their allowance for loan and lease losses to 1.73%, and U. S. banks increased their allowance to 3.12%.

Non-Performing Loans to Total Loans

Non-performing loans are loans that are ninety or more days delinquent in payments of interest and/or principal. In effect, these would be considered bad or toxic assets on the bank's books. On average most banks had tried to maintain a low percentage of non-performing loans to maintain quality and keep earnings up. During 2005, non-performing loans stood at 0.80% of total loans in Texas banks, while U. S banks non-performing loans were at the 0.65% level. For the year of 2006, non-performing loans declined to 0.60% in Texas with U. S. banks also declining to 0.52%. During 2007, non-performing rose slightly to 0.86% in Texas, but U. S. banks increased to 0.87%. With the economic downturn the percentage of non-performing loans to total loans rose substantially in Texas banks to 1.45% in 2008 and 1.84% in U. S. banks. After banks were required to charge-off a substantial number of loans, the level of non-performing loans declined to 0.83% in Texas during 2009, however, the U. S. banks experienced a substantial increase to 3.32% even after substantial charge-offs.

Net Interest Margin

Net interest margin is the difference between the cost funds and amount charged to borrow funds at the bank. A standard considered a good net interest margin is 4.00%. Texas banks did extremely well in their net margins during 2005, averaging 4.27%, while U. S. banks averaged 3.55%. A nominal drop in the net margin occurred in 2006 and 2007 in Texas banks with 4.23% and 4.13% respectively while U. S. banks were dropping to 3.39% and 3.35% respectively. With the downturn, net interest margins dropped substantially in Texas during 2008 to 3.79% and U. S. banks showed a nominal drop to 3.21%. Texas banks improved substantially in 2009 to 4.18% showing economic improvement in the state. U. S. banks improved slightly to 3.50% in 2009.

Percentage of Unprofitable Banks

In good economic times, less that 10 percent of the banks will be unprofitable. In 2005, the number of unprofitable banks in Texas was 5.46% and 6.31% in the U. S. During 2006, the unprofitable rose slightly to 5.76% in Texas and rose to 7.54% in the U. S. As regulators started to see problems in the economy and assessed the banks higher fees during 2007, the unprofitable banks increased to 8.66% in Texas, but increased in U. S. banks to 11.20% seeing some instability nationally. As the economy showed substantial deterioration during 2008, Texas banks jumped to

14.81%% and U. S. banks to 23.38%. Texas banks improved slightly to 13.84% of the banks being unprofitable, while the U. S. unprofitable banks rose to 28.67%.

Percentage of Cease and Desist Orders

Cease and Desist Orders are the most serious actions taken by bank regulators to bring banks back into acceptable performance standards. This measure is a solid measure to determine banks that have serious problems that uncorrected could increase the potential for bank failure. As of March 22, 2010, according to SNL Financial, there were 8,057 financial institutions in the U. S. and 424 of those institutions were under cease and desist orders. This translates into 5 percent of all U. S. banks. Texas had 634 financial institutions with 15 under cease and desist orders, which means that 2 percent of the Texas banks were under these orders.

ANALYSIS

The interrelation of the indicators provide better information on the decline of financial condition of banks than most other measures as noted by Klomp (2010), who performed complex studies of banks in 110 countries, but came to the conclusion that he could not scientifically correlate more than 60 percent of the bank failures which were more related to high economic development. By examining the indicators in relationship to one another, it will give a picture of the interrelationship of the elements to success and failure. Since return on assets has been the key measure of bank performance, which in turn drives return on equity, it will be the control indicator in the analysis with regard to charge-offs, allowance for loan and lease losses, non-performing loans to total loans, and net interest margin. Percentage of unprofitable banks is evaluated individually, since it is only an expression of the overall condition of the banking system.

Return on Assets to Charge-offs

Examining the return on assets to charge-offs both the Texas banks and U. S. banks performed at an acceptable level through 2006. While the Texas banks continued to maintain an acceptable relationship through 2007, the U. S. banks were about three times worse than Texas banks. In 2008, with the downturn, charge-offs in Texas banks increased to 60.5 percent of the return on assets. The U. S. banks in 2008 had a serious deterioration to 101.5 percent of the return on assets. During 2009, Texas banks came back to near acceptable levels at 24.4 percent while U. S. banks went out of control to 2,855.6 percent of return on assets.

Return on Assets to Allowance for Loan and Lease Losses

The allowance in relation to the return on assets for both Texas and U. S. banks performed at an acceptable level through 2007. Beginning with 2008, there was a substantial difference between Texas banks and U. S. banks, with the Texas bank allowance reaching all time levels at 163.0 percent and the U. S. bank level of 1,761.5 percent being 10 plus times higher. Unlike the other measures, in 2009, the regulators demanded more allowance with Texas banks reaching 201.2 percent and the U. S. banks attaining 2,856.0 percent.

Return on Assets to Non-Performing Loans

Similar to the other comparisons, the non-performing loans to return on assets were reasonably consistent through 2007. However, beginning in 2008, there was a great disparity between Texas banks at 179.0 percent non-performing loans in relation to return on assets and U. S. banks 1,415.4 percent or approximately eight times the Texas levels. In 2009, Texas reduced the non-performing loans to return on assets by almost one-half to 96.5 percent. The U. S. banks more than doubled the previous year increasing to 3,688.9 percent.

Return on Assets to Net Interest Margin

Unlike the above elements, with the net interest margin the higher the percentage the better the performance. Both the Texas and U. S. banks performed reasonably consistent through 2007. However, in 2008 the economic downturn caused major separation between the Texas banks at 467.9 percent and the U. S. banks at 103.5 percent. Both U. S. and Texas banks improved in 2009 with Texas banks at a net interest margin percentage of 486.0 percent and U. S. banks at 388.9 percent.

CONCLUSIONS

The analysis comparing the various elements to return on assets shows that, in general, Texas banks performed almost twice as well as U. S. banks in the 2005 to 2009 timeframe. However to examine from a micro standpoint during the most critical period in the economic downturn, it is necessary to compare the performance at December 31, 2008 to December 31, 2009. During this past year, Texas banks return on assets were 7 basis points better in 2009, while U. S. banks were 4 basis points worse. Charge-offs in Texas improved 28 basis points with U. S. banks declining a substantial 125 basis points in the 2008 to 2009 period. With allowance for loan and lease losses indicating increasing potential for losses, the level in Texas banks 51 basis points to 1.73 percent and the U. S. banks only increased 28 basis points but at 2.57 percent was substantial higher. Non-

performing loans in the 2008 to 2009 period in Texas banks was improved by 62 basis points, declining to 0.83 percent. U. S. banks worsened by 184 basis points seriously increasing to 3.32 percent. Both the unprofitable U. S. banks and banks under cease and desist orders are twice or more higher than similar Texas banks.

While both Texas banks and U. S. banks have suffered from the economic downturn, from the data it is established that the Texas banks have performed near twice the level of all the U. S. banks. Additionally, from the data, it is conclusive that Texas banks are improving in 2009, while U. S. banks have not improved as much overall.

REFERENCES

Allen, Franklin & Elena Carletti (2010). An overview of the crisis: causes, consequences, and solutions. *International Review of Finance.* Mar. 10(1), 1-26.

Federal Deposit Insurance Corporation (2010). Quarterly banking profile. Tables I-IIIA for 2005-2009..

Federal Deposit Insurance Corporation (2010). Texas state profile for 2009..

Klomp, Jeroen (2010). Causes of banking crises revisited. *North American Journal of Economics & Finance.* Mar. 21(1), 72-87.

Meadowcroft, John (2010). Three myths of the financial crisis. *Economic Affairs.* 30(1), 107.

SNL Financial (2010). Failed banks: class of 2010. *SNL Data Dispatch*, March 29, 2010. Charlottesville, VA.

EXHIBIT 1: KEY BANKING ELEMENTS					
	2005	2006	2007	2008	2009
Return on Assets					
Texas Banks	1.32%	1.33%	1.24%	0.81%	0.86%
U. S. Banks	1.30%	1.33%	0.93%	0.13%	0.09%
Charge-offs					
Texas Banks	0.23%	0.19%	0.22%	0.49%	0.21%
U. S. Banks	0.56%	0.41%	0.62%	1.32%	2.57%
Allowance for Loan & Lease Losses					
Texas Banks	1.19%	1.11%	1.09%	1.32%	1.73%
U. S. Banks	1.12%	1.15%	1.35%	2.29%	2.57%
Non-Performing Loans to Total Loans					
Texas Banks	0.80%	0.60%	0.86%	1.45%	0.83%
U. S. Banks	0.65%	0.52%	0.87%	1.84%	3.32%
Net Interest Margin					
Texas Banks	4.27%	4.23%	4.13%	3.79%	4.18%
U. S. Banks	3.55%	3.39%	3.35%	3.21%	3.50%
Percentage of Unprofitable Banks					
Texas Banks	5.46%	5.76%	8.66%	14.81%	13.84%
U. S. Banks	6.31%	7.54%	11.20%	23.38%	28.67%

DETERMINANTS OF VALUE AND PRODUCTIVITY IN A COMPLEX LABOR MARKET: HOW SABERMETRICS AND STATISTICAL INNOVATION CHANGED THE BUSINESS OF PROFESSIONAL BASEBALL

Brent C. Estes, Sam Houston State University
N. Anna Shaheen, Sam Houston State University

ABSTRACT

Professional baseball as an industry mirrors many organizations in today's business world in terms of its need to objectively evaluate the performance of its workers (players). Baseball relies on these evaluations in order to establish essential aspects of the game such as strategizing, scouting talent, drafting amateur players, negotiating, signing/resigning free-agents, calling-up minor leaguers, trading players, and releasing players. In addition, owners and team executives are constantly trying to answer the same fundamental questions: Are we getting the production we are paying for? Does player performance decline with increased job security? To what extent does money motivate players? What is a player's replacement value? In professional baseball, the performance of a player varies from game to game and from season to season. Due to this randomness of productivity, it is impossible to absolutely know the value of a player's inputs relative to his outcomes. Therefore, a player's productivity as it relates to determinants of value must be assessed by using reliable measurements of performance indicative of his expected contributions. Given the current landscape of baseball's labor market, it is especially important for team owners and executives to be able to determine, with some degree of certainty, a player's performance value. With skyrocketing player salaries and the ever-diminishing realization of competitive balance, the success of an organization hinges on its ability to make correct personnel decisions in terms signing and resigning players. This study examines two different methods of assessing Major League Baseball player performance as it relates to evaluating productivity, and illustrates how statistical innovation is changing 165 years of traditional baseball wisdom and ultimately, the business of professional baseball.

INTRODUCTION

Accurate assessment of employee performance and productivity is an invaluable part of understanding, predicting, and influencing organizational success. Likewise, evaluation of worker development in terms of skill, ability, and accomplishments is essential in determining an individual's value to a company and its components. The knowledge and insight gained from these assessments allows organizations to become more efficient and ultimately more effective. In today's business world, many companies rely on employee performance assessments to determine salaries and rewards for their workers.

Bishop (1987) noted that adjusting salaries to reflect productivity produces three kinds of benefits for an organization. First, it serves as an incentive for greater effort from the employee. Second, it tends to attract more productive workers who like to work hard. Third, it reduces the probability of losing the best performers to other companies and raises the probability that the least productive workers will leave.

In most cases, however, performance cannot be measured objectively because there is no universal standard. What one employer may value and consider productive, another employer may regard as insufficient and lacking. According to Alchain and Demsetz (1962), this problem is a fundamental contributor to an organization's inability to accurately measure employee productivity, especially in regards to long-term labor contracts. Many organizations are left trying to answer the same age old questions: Is employee "A" as productive as in years past? How does employee "A" compare to employee "B"? Could I replace employee "A's" production value with employee "C" at a lower cost to the organization?

These questions and others like them continue to present many organizations with legitimate challenges in their attempts to assess employee performance and determine how it translates into value for their company. According to Pinder (1984), employee performance is often difficult to ascertain and predict due in large part to the subjective evaluations used to measure performance. While Pinder's assertion is accurate for numerous organizations throughout many industries, it does not hold true for work environments where performance is not measured subjectively, such as professional sports, in particular, professional baseball. Professional baseball, as an industry, is unique in that workers (players) can be evaluated by the same impartial performance standards. These objectively measured standards are easily quantifiable and are capable of being compared interchangeably with those of past generations.

Methods of evaluation in baseball rely on statistical measures of individual and team performance. The standard for most measurements in baseball is perfection. Nearly every percentage statistic in baseball is a number signifying a proximity to perfection.

The use of statistics in baseball, and sports for that matter, is not a new concept. As the game of baseball has evolved, however, so has the complexity of its statistical measures. Likewise, the way that players are evaluated and statistics are analyzed has changed dramatically during the past

half century. Over time, more accurate, detailed, descriptive and efficient ways to measure talent, performance, and productivity have been developed. Specifically, in the last thirty years objective measurements called sabermetrics have redefined statistical analysis and generated a buzz throughout the baseball world (Berardino, 2003).

THE LABOR MARKET OF PROFESSIONAL BASEBALL

In professional baseball, the performance of a player varies from game to game and from season to season. Due to this randomness of productivity, it is impossible to absolutely know the value of a player's inputs relative to his outcomes. Rather, a player's productivity must be estimated from reliable measurements used to determine his expected contributions (Krautmann, 1990). In addition, since past performance is the primary tool used to assess future productivity, it is imperative for evaluators to understand why players performed better or worse in certain years and what factors contributed to their improvement or decline in production.

Given the current landscape of baseball's labor market, it is especially important for team owners and executives to be able to determine, with some degree of certainty, a player's performance value. With skyrocketing player salaries and the ever-diminishing realization of competitive balance, the success of an organization hinges on its ability to make correct personnel decisions in terms of signing and resigning players.

In today's baseball labor market, teams simply do not have the financial resources to sign every player they want to a contract (although the New York Yankees have attempted to prove otherwise). Teams must allocate their monetary resources to available player talent without compromising their budgetary limitations. As such, an organization's objective should be to use all available information to make decisions that will maximize the team's probability of winning games (Hadley, et. al, 2000).

Professional baseball as an industry mirrors many organizations in today's business world in terms of its need to objectively evaluate the performance of its workers (players). Baseball relies on these evaluations in order to establish essential aspects of the game such as strategizing, scouting talent, drafting amateur players, negotiating, signing/resigning free-agents, calling-up minor leaguers, trading players, and releasing players. In addition, owners and team executives are constantly trying to answer the same fundamental questions: Are we getting the production we are paying for? Does player performance decline with increased job security? To what extent does money motivate players? What is a player's replacement value?

The definition of success for a baseball general manager is to be able to accurately answer the preceding questions. As a result, much research and painstaking effort has gone into finding the best approach to objectively analyze the game, determine player value, and address problems associated within the industry. For nearly a century, conventional wisdom with respect to traditional objective measures of player performance superseded the fundamental calling for a paradigm shift

among passionate and otherwise intelligent observers of the game. However, around the beginning of the second half of the 20th Century, alternative ways to objectively measure success and failure among players began to emerge. Despite heavy resistance that continues to permeate throughout the game today, these alternative metrics, known in the modern baseball era as sabermetrics, have found a foothold among both experts and novices alike.

SABERMETRICS

According to Bill James, well-known author, sabermetrician, and baseball theorist, sabermetrics is "the search for objective knowledge about baseball." The term is derived from the acronym SABR, which stands for the *Society for American Baseball Research* (James, 1982). It was coined by James, a Kansas baseball fanatic whose self-published *Baseball Abstracts* in the 1970s and 1980s brought sophisticated mathematical tools to the masses for the first time. James, who is widely considered to be the father of sabermetrics, began writing his unorthodox and original essays on the game of baseball in 1977 (deMause, 2002).

The idea behind sabermetrics is to find a way to objectively analyze every aspect of the game. Using sabermetrics means relying on probabilities and scientific standards instead of the naked eye, no matter how much that eye has seen before (Quinn, 2003). Sabermetrics offers more comprehensive and complete assessments of performance than traditional measures.

These "new" statistical metrics and their utilization have become an integral part of professional baseball. Throughout the game, sabermetrics is used to develop strategy and assess team strengths and weaknesses. Yet, sabermetrics is primarily concerned with determining players' past and present values and predicting their future performance. Thus, sabermetrics is often used by teams and agents alike to evaluate a player's performance in relation to other players for the purposes of negotiating contracts.

According to Frank Coonelly, current President of the Pittsburg Pirates and former Senior Vice President for Labor Relations for Major League Baseball, sabermetrics has become the language of salary arbitration and salary negotiation. Coonelly said:

> "There used to be the argument that (the classic statistics) were the only official evidence. The union felt the clubs had better access to the 'exotic' statistics than they did. All of that went by the wayside, probably 10 years ago or less, when STATS, Inc. came out with their handbook. Immediately, everybody in the arbitration room would have the handbook" (Quinn, 2003).

Incidentally, STATS, Inc. was created by Bill James and his cohorts, John Dewan and Dick Cramer in 1988, in an effort to establish a pitch-by-pitch, play-by-play database for every game

played during the season. It has since evolved into a mainstay of professional baseball relied on by the media and a number of front offices throughout baseball.

Although originally only appealing to hardcore statisticians with a passion for baseball, James's ideas have since made their way to mainstream outlets and spawned a new generation of statistical gurus determined to change the dynamics of how the game is analyzed. Lehman (1984) stated that sabermetricians "have sparked a mini-revolution as startling in its way as the adoption of the designated hitter rule by the American League a decade ago" (p. 75). It would take another decade for baseball to recognize the "mini-revolution" that Lehman (1984) referred to. Nevertheless, it happened just the same.

Among James's disciples are Billy Beane, the Oakland Athletics' General Manager and "sabermetrician extraordinaire" on whom the bestselling book *Moneyball* is based; Beane's mentor and predecessor, Sandy Alderson, who is the former General Manager of the Oakland Athletics and Chief Executive Officer for the San Diego Padres; Beane's one time assistants J. P. Ricciardi and Paul DePodesta, former general managers for the Toronto Blue Jays and Los Angeles Dodgers respectively (deMause, 2002). Other notables working with sabermetrics include Beane's close friend, Kevin Towers, former General Manager for the San Diego Padres and current Special Assistant to the General Manager of the New York Yankees and Towers's one time assistant, Theo Epstein, the 37-year-old General Manager of Boston Red Sox who was hand picked by new owner, John Henry to challenge decades of baseball wisdom by basing important decisions in large part on objective research, or what baseball's new generation calls "sabermetrics" (Birger, 2003).

As the once youngest GM in baseball history, Theo Epstein's experience as a ball player is limited to his days playing for the local high school team in Brookline, Massachusetts ("Meet Boston GM," 2002). Further, he had limited management experience and lacks a true baseball pedigree. In many baseball circles, a person possessing these qualifications or lack there of would be considered highly under-qualified with regards to such a high profile job- the general manager of one of the most storied franchises in sports history. However, many see the hiring of Epstein and others like him as an indication of things to come. Incidentally, during his brief tenure, Epstein engineered the first World Series championship by the Red Sox in 86 years in just his third season as General Manager in 2004 and a second world championship in 2007.

"I think what we're seeing is the beginning of something much bigger," predicts ESPN.com columnist Rob Neyer. "In five to ten years at the most, half of the GMs in baseball will have this sort of background [speaking in reference to sabermetrics]." Even Bill James himself has joined the professional ranks. He was hired by Epstein and the Red Sox in 2003, as the club's senior advisor on personnel matters (deMause, 2002).

Pioneer sabermetrician Craig Wright, in his forward for Bill James's *1985 Baseball Abstract,* describes the basic concept of sabermetrics. He explains:

> "Sabermetrics is the scientific research of the available evidence to identify, study, and measure forces in professional baseball. A sabermetrician is not a statistician. Sabermetricians do *not* study baseball statistics. Sabermetricians are actually involved in research, scientific study, and the subject is baseball. The real tools of the trade fall under scientific methodology. Besides statistical techniques and applications it [sabermetrics] includes things like rules of evidence, rules of logic, testing theories and measures by internal consistency, relation to known quantities and qualities, and common sense." (p. 1)

While it is true that sabermetricians are not necessarily statisticians, a great deal of sabermetrics involves understanding how to use statistics properly and deciphering which statistics are useful for what purposes. Since statistics are often the best objective record of the game available, sabermetricians often use them in their attempt to answer objective questions about baseball, such as "which player on the Astros contributed the most to the team's offense?" or "How many home runs will Albert Pujols hit next year?" Sabermetrics cannot logically deal with subjective judgments, such as "Who is your favorite team?" or "George Steinbrenner is bad for the game of baseball" (Grabiner, *The Sabermetric Manifesto*). Sabermetrics can provide a more objective, comprehensive measure of player performance.

As the former General Manager for the Oakland Athletics, Sandy Alderson was the first general manager in baseball to adopt the sabermetric philosophy. Mr. Alderson not only believed in this alternative approach, he built championship teams around its ideology. Like Bill James, Sandy Alderson is considered a pioneer in terms of sabermetrics and its influence in the game of baseball today. Mr. Alderson relied on sabermetrics, in his words, "quite extensively" in order to gain an edge over the competition. Alderson describes what attracted him to the sabermetric philosophy:

> "What was in my favor was the fact that I was new. I came into baseball from an entirely different world. I wasn't bound by any sort of tradition or experience or wisdom that I had received. It's not something I was burdened by. So I started looking around and thinking more independently and critically and it was about that time that these studies were reported and discussed. They seemed appealing to me from the standpoint of objectivity…… What struck me, what got my attention way back when, was that certain statistics could be tied directly to outcomes. Through a regression analysis, basically, you could determine which variables were most important in reaching a particular conclusion and if you could show, for example, that teams who have the highest run differential between what they score and what they give up tend to have the best winning percentages, and you start working from the proposition that you want to give up as few runs over the course of a season and

you want to score as many as you can, you start looking at the probabilities of reaching those two results. In order to maximize the differential, you come up with variables that lead you to on base percentage, power, not walking anybody from a pitching standpoint. Then what you end up doing is trying to identify the players that give you the best potential and you start emphasizing certain things over others. You begin to prioritize things and rely on certain things that seem to be more indicative and predictive of the end result, winning. Batting average is not predictive of anything" (S. Alderson, personal communication, September 1, 2005).

CHALLENGING CONVENTIONAL WISDOM

The very idea of sabermetrics contradicts over a century of traditional baseball wisdom. For one to suggest that there are better ways to analyze "America's Pastime" is considered by many as, in no uncertain terms, blasphemy. However, this notion to challenge conventional analysis did not begin with "sabermetrics" as we now know it. In fact, on August 2, 1954, an article appeared in *LIFE* magazine written by Hall of Fame Executive, Branch Rickey. In the article, Rickey, recognized by many as one of the greatest baseball management minds of all-time, mentioned the development of a new way to analyze the game of baseball based on his own examination of performance standards and their value to winning and losing games, a clear precursor to sabermetric theories and principles. Specifically, Mr. Rickey, with the help of mathematicians from M.I.T., set forth a formula that would predict how many games a team would win based on various commonly available team statistics. Rickey writes:

"Baseball people generally are allergic to new ideas. We are slow to change. For 51 years I have judged baseball by personal observation, by considered opinion, and by accepted statistical methods. But recently I have come upon a device for measuring baseball which has compelled me to put different values on some of my oldest and most cherished theories. It reveals some new and startling truths about the nature of the game. It is a means of gauging with a high degree of accuracy important factors which contribute to winning and losing baseball games. It is the most disconcerting and at the same time the most constructive thing to come into baseball in my memory…..If the baseball world is to accept this new system of analyzing the game- and eventually it will- it must first give up preconceived ideas. I had to. The formula outrages certain standards that experienced baseball people have sworn by all their lives. Runs batted in? A misleading figure. Strikeouts? I always rated them highly as a determining force in pitching. I do now. But new facets convince me that I have overrated their importance in so far as game importance is concerned. Even batting average must be reexamined….." (p.78).

Branch Rickey used his new theories to rebuild a struggling Pittsburgh franchise that had lost at least 90 games in each of the previous four years. In 1955, though, the team improved in almost every statistical category and the Pirates' winning percentage climbed over .400. By 1958, the team was over the .500 mark for the first time in nearly a decade. Yet, the culmination of Branch Rickey's rebuilding efforts occurred in 1960, when the Pirates, led by a core of talent developed by Rickey, won the World Series over the New York Yankees (Woolner, 1997).

Sabermetrics is not a phenomenon that emerged with the arrival of Bill James and personal computers. Even before Rickey's article, scholars began inundating academic journals with sophisticated analyses of baseball. In 1952, Harvard statistics professor, Frederick Mosteller used binomial probability theory to prove that the best-of seven World Series was an inadequate and unreliable format to determine baseball's champion. In 1956, an article in *American Statistician* proposed a method to adjust league standings based on a team's strength of schedule. Four years later, a paper was presented to the American Statistical Association titled "The Distribution of Runs in the Game of Baseball," which was the first advanced attempt to combine the probabilities of hits, walks, outs, and more into a model of how runs score (Schwarz, 2004).

In the early 1960s, a Johns Hopkins professor named Earnshaw Cook began compiling significant amounts of data that would overturn baseball's conventional wisdom. Cook then presented his findings to executives for a handful of struggling teams. Cook was largely ignored, so, in 1964, he wrote a book titled *Percentage Baseball* (Surowiecki, 2002). In the book, Cook's theories used stochastic analysis to derive performance criteria for both teams and individual players that were reasonably successful absolute measures. Many sabermetricians consider Cook's book to be the original sabermetric manuscript and the foundation for much of the baseball research that we have today. And while nearly five decades have passed since Cook first published his findings, it was not until recently that the baseball world began to embrace sabermetrics (Schwarz, 2004).

THE APPLICATION OF SABERMETRICS

Over the past 30 years, Bill James's work on player evaluation, player development, and baseball strategy has gone largely unnoticed. While James had a dedicated following of readers, many of whom went on to expand James's work doing ground breaking statistical analysis of their own, most baseball owners and general managers simply ignored him. In the past ten years, however, all of this has changed. The new acceptance and recognition of sabermetrics can be attributed directly to the success of the Oakland Athletics, who, thanks in no small part to General Manager, Billy Beane's clever application of sabermetric insights, brought James new attention (Surowiecki, 2003). Several baseball executives had tinkered with the sabermetric method in the past. However, Beane was the first general manager to build his organization around sabermetrics (Surowiecki, 2002). Beane's extraordinary success is chronicled in Michael Lewis's bestselling book, *Moneyball: The Art of Winning an Unfair Game.*

Lewis (2003) focuses on the phenomenal accomplishments of Beane, who has produced great teams despite one of the lower payrolls in baseball. Since taking over as general manager in 1999, the Athletics have compiled a remarkable record. Consider the numbers. In 1999, Oakland ranked eleventh (out of fourteen teams) in the American League in payroll and fifth in wins. In 2000, the Athletics ranked twelfth in payroll and second in wins, a feat they duplicated in 2001. In 2002, they ranked twelfth in payroll again, and first in wins (Thaler & Sunstein, 2003).

The foundation for Lewis's book is based on the acceptance of baseball's ever-changing economic landscape. Since the inception of free-agency, market demands in terms of higher salaries and longer contracts have drastically increased- allowing only the wealthier teams to contend for elite talent. In turn, this has created significant gaps between larger and smaller market teams with respect to competitive balance. Without salary cap restrictions, large market owners are able to "stockpile" premier players, leaving small market owners with fewer resources with which to build contending teams (Lewis, 2003).

Ultimately, according to Lewis (2003), a small market team's success is contingent on the general manager's ability to identify undervalued, overachieving talent. This new requisite for producing competitive ball clubs led Oakland's Billy Beane to the work of Bill James. As an assistant general manager under Sandy Alderson, Beane was indoctrinated into an alternative approach to evaluate performance void of subjective judgment regarding a player's potential and his "intangibles." With regards to the evaluation of players, Alderson said:

> "Clearly, along some point, potential has to convert into performance andsomewhere along the line I think it's less worth while to rely on the potentiality of a player, and it becomes more realistic and more relevant to rely on performance of that player. When I talk about potential, I talk about the raw potential: somebody's speed, somebody's power, somebody's throwing arm, all of the things that in combination can lead to a successful player. At some point you have to be less indirect in your analysis. You look at what the player has done, look and see whether or not that is predictive of what the player will do in the future. The whole business of sabermetrics is, first and foremost, adopting statistics rather than subjective evaluation, and second, it's finding out which statistics are most relevant to that analytical approach" (S. Alderson, personal communication, September 1, 2005).

Throughout *Moneyball*, Lewis (2003) outlines Beane's unconventional strategies for success, which are consistent with fundamental sabermetric theories and ideas created by pioneering sabermetricians such as Bill James, Craig Wright, John Thorn, and Pete Palmer. In addition, the author illustrates how these principles, adopted by Beane, changed the way that players were evaluated. By relying on objective statistical analysis, rather than instinct and subjective measurements, Beane was able to defy traditionalist baseball mentality and create a competitive

team with limited financial resources. As a result, the baseball world took notice. While there is still an apparent loyalty to conventional baseball performance measures, few can argue with Beane's accomplishments.

The success of the Oakland A's has sparked some baseball insiders to reevaluate the use of statistics in analyzing performance. Up until the late 1990s, evaluating baseball talent and player performance had consisted of relying on misleading measurements of things like speed, power, hitting ability, and arm strength (Lewis, 2003). According to Quinn (2003), it's simply a matter of differing philosophies. It's the statistical methods of evaluation versus the time-honored strategies of experts who have scouted, played, and breathed baseball for decades (Thaler & Sunstein, 2003).

For example, the old guard says sign players with inherent athleticism. Ignore performance numbers. Trust gut instincts and the eye of experience. Tools are what matter most. On the other hand, the new guard says numbers- objective numbers- tell the true story, and that performance is more important than raw talent (Quinn, 2003). So, what is the verdict? If Billy Beane's success is any indication, then statistical methods will outperform the experts more often than not.
According to Quinn (2003), the idea behind sabermetrics is not just using certain prescribed methods to analyze baseball. Rather, the real purpose is to find a way to objectively analyze every facet of the game. Applying sabermetric principles means relying on probabilities and scientific standards instead of intuition and experience.

Since sabermetrics is primarily concerned with determining the value of a player, one of the most common applications of sabermetrics is the evaluation of offensive performance. According to James (1984), a team's offense is comprised of two parts: the ability to get players on base while avoiding outs and the ability to advance runners. There are various ways to measure offensive performance and several levels of complexity for different evaluation methods, yet all of them rely on measuring those two facets of offense: On-Base and Advancement.

TRADITIONAL STATISTICS VERSUS SABERMETRICS

One of the more traditional measures of offensive performance is batting average. A player's batting average, once baseball's gold standard of hitting ability, is considered by many sabermetricians to be a statistic of limited usefulness because it has been proven to be a poor predictor of a team's ability to score runs (Thorn & Palmer, 1990). Batting average really only measures a player's ability to hit, and while batting titles are awarded to players with the best average, victories go to the teams with the most runs (Quinn, 2003).

Thorn and Palmer (1985), in their book *The Hidden Game of Baseball*, argue against using traditional performance measures such as batting average, runs batted in, home run totals, and runs scored to evaluate a player's worth due to their extreme unreliability and their likelihood to be misinterpreted. With respect to batting average, Thorn and Palmer (1985) write:

> "The batting average remains the most hallowed statistic of baseball, despite its shortcomings: it makes no distinction between a bunt single and a home run, gives no indication of the effect of each hit, and fails to account for bases reached by walks, errors, and hit batsmen……A two out bunt single in the ninth inning with no one on base and your team trailing by 6 runs counts the same as Bobby Thompson's "shot heard 'round the world"; and no credit for fouling off 7 strikes after gaining a full count to earn a walk is given in the batting average" (pp. 17, 23).

Barry (1988) mentions that by relying on batting average alone to determine a player's performance level devalues the accomplishments of the extra-base hitters, players who draw walks, and clutch hitters. In addition, James (1985) argues that batting average is an over-weighted offensive statistic that is limited in its interpretive value.

Another traditional offensive performance measure with apparent limitations is runs batted in (RBI). RBI is an incomplete measure used to evaluate hitters that is situation dependent based on opportunities out of the batter's control. For example, the amount of runs that a player bats in depends largely on where he hits in the lineup and entirely on the number of runners on base (Thorn and Palmer, 1985). In the same way, hitting a homerun with the bases empty counts for one RBI, yet hitting a homerun with bases loaded counts for four RBI. The individual contribution of the hitter does not change. However, the difference is entirely dependent on the hitter's teammates' ability to get on base. Therefore, using RBI to evaluate individual hitters is problematic (Huckabay, 2003).

Thorn and Palmer (1985) point out other baseball performance measures that are either flawed or situation dependent. Specifically, they mention the following: (1) Stolen Bases- the amount of stolen bases a player has is not indicative of his base running ability; the player may have been caught stealing as often as he stole, costing his team runs (p. 27); (2) Slugging Percentage- a player's slugging percentage can be improved by a bunt single, which is not a measure of "slugging" ability (p. 24); (3) On-Base Percentage- OBP makes no distinction between a walk and a grand-slam home run (p. 25); (4) Earned Run Average- a pitcher's ERA fails to penalize a player who "retires the first two batters, watches a ground ball get booted by his shortstop, and then yields 6 home runs" (p. 29); Win-Loss Records- a pitcher's win-loss record is entirely dependent on the number of runs his team scores, thus making it an inaccurate measure of actual pitching performance (p. 28); Saves- a pitcher can earn a save without actually retiring a batter (p. 33); Fielding Percentage- this performance measure does not factor in a fielder's range (a fielder can not make an error on a ball he does not touch) (p. 33).

Thorn and Palmer (1985) maintain that many baseball statistics used to evaluate players are often misleading, inaccurate, and incomplete measures of performance. They note that baseball traditionalists rely far too heavily on a one-dimensional approach to evaluate a player's production and contribution to his team, oftentimes ignoring logic in favor of core conventional methods. In addition, most of today's performance measures are only meant to reveal parts of a player's

production value. According to James (1984), (1) a "clean" measure of performance is always to be preferred to a "situation dependent" measure and (2) an accurate measure of performance is always preferred to a less accurate measure; hence the creation of sabermetrics.

Sabermetricians, such as Bill James (1984), argue that performance should be analyzed by using multi-dimensional measures that can be utilized and interpreted in terms of a ballplayer's purpose for playing baseball: to do things which create wins for his team, while avoiding those things which create losses for his team. In other words, since an offensive player's job is to create runs for his team, then a hitter's performance should be measured in terms of his ability to generate runs. Likewise, since a defensive player's job is to avoid giving up runs, then a fielder's and a pitcher's performance should be measured according to his ability to prevent the opposing team from scoring (James, 1984).

The point of sabermetrics is to make baseball statistics more explicable, not less, by reducing performance to a set of easily quantifiable "metrics" (deMause, 2002). Some of the more popular and well known metrics developed by sabermetricians are On-Base Percentage plus Slugging Average, Total Average, Runs Created, Total Offensive Production Rating, Total Pitching Effectiveness Rating, Win Shares, Total Player Rating, Major League Equivalency, Pythagorean Method, Range Factor, and Walks plus Hits per Innings Pitched.

Perhaps the most recognized and widely accepted sabermetric statistic is OPS, which stands for On-Base Percentage plus Slugging Average. OPS is frequently used by sports writers and journalists and is often referred to by sports broadcasters on television programs such ESPN's *Baseball Tonight.* According to deMause (2002), OPS has caught on with the baseball world because it is easy to calculate and it is an excellent predictor of runs scored. Specifically, OPS credits hitters with getting on base and advancing runners. OPS is calculated according to the following equation:

$$OPS = (H+BB+HBP)/(AB+BB+HBP+SF) + (H+2B+(2*3B)+(3*HR))/AB$$

H = hits
BB = walks
HBP = hit by pitch
SF = sacrifice fly
2B = doubles
3B = triples
HR = home runs
AB = at bats

Another popular sabermetric statistic is Total Average. Invented by baseball writer, Tom Boswell in the 1980s, Total Average measures a baseball player's offensive contribution from a

variety of batting and base running events. It is determined by calculating the total number of bases that a player earns divided by the total number of outs that a player produces. Boswell (1985) explains:

> "Take Tim Raines as an example. The Expo outfielder had 137 singles, 38 doubles, 9 triples, 8 home runs, 87 walks, 2 hit-by-pitches, and 75 stolen bases. Subtract 10 bases for the 10 times he was caught stealing, leaving 426 bases. Raines also came to bat 622 times and got 192 hits, which meant that the other 430 times he made an out. Add to this the 10 times he got thrown out stealing, plus an extra out for each of the 7 times he grounded into a double play. That makes 447 outs. Now divide the bases by the outs and you get Total Average- .953 for Raines, the best in the National League in 1984" (p. 27).

Also intended to measure offensive productivity, Runs Created was developed by Bill James in 1979, in an attempt to estimate the number of runs that a batter creates for his team. Runs Created accounts for offensive productivity per plate appearance and playing time. According to James (1979), a hitter's success should be measured in terms of what he is trying to do, create runs. James (1979) in his *Baseball Abstract* wrote:

> "I find it remarkable that, in listing offenses, the league will list first-meaning best-not the team which scored the most runs, but the team with the highest batting average. It should be obvious that the purpose of an offense is not to compile a high batting average" (p. 23).

In response to what James called a "real need" in the statistical landscape of the game, he set out to develop a formula that takes the numbers of hits, walks, doubles, triples, home runs, and other offensive contributions and express them all as runs (James, 1984). Runs Created can be calculated by the following formula:

$$RC = \frac{(H+BB+HBP-CS-GIDP)(TB+.26(BB-IBB+HBP)+.52(SH+SF+SB))}{AB+BB+HBP+SH+SF}$$

H = hits
BB = walks
HBP = hit by pitch
SF = sacrifice fly
AB = at bats
CS = caught stealing

GIDP = grounded into double play
TB = total bases
IBB = intentional walk
SH = sacrifice hits
SB = stolen bases

The Runs Created metric is highly correlated with a team's total run production. In other words, one could plug actual numbers from past seasons into the equation and determine the number of runs a team scored for any given year. In fact, Runs Created has been proven to equal the actual number of runs a team scores in a season to within 5% (James, 1984). In 2003, the Atlanta Braves led the National League in runs scored. They also had the most runs created. That same year, the St. Louis Cardinals, Colorado Rockies, and Houston Astros were second, third, and forth respectively in each category.

James also invented the Pythagorean Formula which can predict a team's winning percentage by taking its runs scored squared and dividing by the sum of its runs scored squared and its runs allowed squared. The concept is based on what James (1984) calls one of the "known principles of sabermetrics": there is a predictable relationship between the number of runs a team scores, the number they allow, and the number of games they will win. Empirically, the Pythagorean Formula correlates fairly well with how teams actually perform (Thorn & Palmer, 1990). In fact, in his book *Moneyball*, Michael Lewis (2003) describes how Paul DePodesta, the then assistant general manager of the Oakland Athletics, used James's formula to predict, with significant accuracy, how many runs the A's would need to score for the 2002 season in order to make the playoffs. Lewis writes:

> "Before the 2002 season, Paul DePodesta had reduced the coming six months to a math problem. He judged how many wins it would take to make the playoffs: 95. He then calculated how many more runs the Oakland A's would need to score than they allowed to win 95 games: 135. (The idea that there was a stable relationship between season run totals and season wins was another Jamesean discovery.) Then, using the A's players' past performance as a guide, he made reasoned arguments about how many runs they would actually score and allow. If they didn't suffer an abnormally large number of injuries, he said, the team would score between 800 and 820 runs and give up between 650 and 670 runs. From that he predicted the team would win between 93 and 97 games and probably wind up in the playoffs." (The A's wound up scoring 800 and allowing 653) (p. 124).

Two metrics designed to evaluate hitting and pitching production are Total Offensive Production Rating (TOPR) and Total Pitching Effectiveness Rating (TPER). Developed by Hitzges and Lawson in 1994, TOPR and TPER account for a player's total production performance based

on total bases earned and lost. In addition, TOPR and TPER are tools that determine both hitting and pitching performances by using a single measurement. In other words, TOPR and TPER allows offensive and pitching production to be compared by the same performance standard based on each player's role: a hitter's role is to produce bases and minimize outs, while a pitcher's role is to generate outs and minimize bases (Hitzges and Lawson, 1994).

In the same way, Pete Palmer's Total Player Rating is a metric used for measuring a baseball player's production value. This particular measurement allows players to be compared against each other from different teams, different leagues, and across different eras. The concept is based on assigning run values to various aspects of hitting, pitching, and fielding performance using linear weights.

Win Shares, developed by Bill James in 2002, is another metric that evaluates players based on complete performance measures. Specifically, Win Shares assigns players fractions of their team's wins based on individual hitting, pitching, and fielding performance. Win shares differs from other player rating metrics in that it is based on team wins, not runs. Win Shares is an exhaustive statistic that sums up player's contribution to his team in a single number.

According to James (2002), Win Shares was invented as a simplistic way to compare players; players from different positions, players from different teams, and players from different eras. In his most recent book of the same name, *Win Shares,* Bill James explains his rationale for developing his newest sabermetric measurement. James writes:

> "For many years, I have wanted to have a system to summarize each player's value each season into a simple integer. Willie Mays' value in 1954 is 40, in 1955, 40, in 1956, 27, while Mickey Mantle in the same three years is 36, 41, 49. If we had an analytical system *in which we had confidence*, and which delivered results in that simple a form, it would open the door to researching thousands of questions which are virtually inaccessible without such a method. It would reduce enormously the time and effort required to research other questions, which can be accessed by other methods, but only with great difficulty." (p. 3).

While the actual Win Shares methodology is somewhat complex, the results are groundbreaking (Neyer, 2002). The formula itself credits players proportionally based on their statistics. A Win Share is actually the number of wins contributed by a player multiplied by three. Conversely, the formula credits a team with three win shares for each win. For example, if a team wins 100 games in a season, the players on that team are credited with 300 Win Shares (James, 2002). According to James (2002), the three to one ratio is important in order to provide a meaningful distinction between players.

There are three types of Win Shares: hitting, fielding, and pitching. In general, hitting contributions receive 48% of the Win Shares, 35% are assigned to pitchers, and 17% are assigned

to fielders. However, those percentages can vary based on individual team strengths and weaknesses (James, 2002). In his book *Win Shares*, James introduces a short-form method for calculating Win Shares as well as a long method. The short form method is represented by the following equation:

$$WS = \frac{RC - (Outs/12)}{3}$$

RC = Runs Created

The long method of calculating Win Shares is a tedious and intensive process. The long method is based on identifying what James refers to as "marginal runs." Marginal runs are all runs scored by a team minus one-half the league average and all runs allowed by a team less than one and one-half times the league average. In addition, this method involves determining the ratio of Win Shares credited to the offense and defense, which is based on park-adjusted runs scored and allowed. Then, Runs Created are calculated as well as outs made by each hitter. "Claim points" are used to divide up offensive and defensive win shares. Finally, individual win shares are determined for each player on the team.

In an effort to translate minor league data into major league performance, Bill James created a metric called MLE which stands for Major League Equivalency. By adjusting for the run environment, caliber of competition, and park factors, James was able to estimate how minor league players would perform at the major league level given the same production. MLEs are not a prediction of what a player will do, just a translation of what the major league equivalence of what the player actually did. However, MLEs, like major league statistics, have strong predictive value (James, 1985). In reference to MLEs, James (1989) wrote, "In my opinion, this is the most important thing I've learned in my years of studying sabermetrics in terms of its potential ability to help a baseball team" (p. 475).

Most sabermetric statistics are objective offensive measures of performance. A few metrics do, however, measure some of the defensive aspects of the game. An example is Range Factor. Range Factor is a metric created to quantify a player's fielding ability, beyond just errors. It is calculated by multiplying assists and putouts by nine, then dividing by the number of innings played (Quinn, 2003).

Another metric used to measure non-offensive performance is WHIP, which stands for walks plus hits per innings pitched. WHIP is a sabermetric tool used to measure a pitcher's ability to prevent hitters from getting on base. WHIP is another "mainstream" sabermetric statistic that is widely used and recognized throughout baseball. Most newspaper box scores now print a pitcher's WHIP in addition to earned run average and strikeouts.

What is different about today? Does the difference between the old statistics and the new statistics matter? T. J. Quinn, a sports writer for the *New York Daily News*, has an interesting answer. Quinn (2003) writes:

> "It's like physics. In the late 17th century, Isaac Newton laid out his laws of gravity, and for the most part those rules work for the average person today. Then, in the early 20th century, Albert Einstein came along and proved that Newton's ideas about gravity missed the point (the point being relativity). But it was the sort of difference that mattered only to scientists. If you want to know why space and time bend, it matters. If you're someone who wants to drop a water balloon off a building, it doesn't" (p. 3).

According to Gillette (1993), the familiar traditional measures of player performance- batting average, home runs, runs batted in, wins, losses, earned run average, and strikeouts- have merit, or they wouldn't be universally known throughout baseball. However, Gillette argues that the traditional statistics have not evolved with the game of baseball. Gillette writes:

> "What's wrong with the old familiar stats is that they haven't changed as the game has changed. They still have value, but newer stats are needed to describe and analyze the way the game is played today. Everything else in life changes over time, so why should baseball- or baseball statistics, for that matter- remain frozen? ……Scholars, doctors, lawyers, politicians, teachers, stock brokers, mechanics, and many others play largely the same roles in society today as they did 75 years ago. Yet their training isn't the same, their tools are different, the way they approach their jobs has changed, and the amount they get paid has increased manifold. There is no reason to assume that similar changes haven't affected our national pastime. Just as our language has altered and our sciences have progressed, so, too, must our understanding of how baseball is played. Baseball statistics are the measure of the game, and the 'new' statistics are simply an attempt to assess the modern game more accurately" (p. v).

Bill James (1989) wrote, "The evolution of statistical information about baseball, progressing nicely from about 1869 to 1955, was frozen solid for a generation afterward" (p. 453). However, with the emergence of sabermetrics, statistical analysis of baseball performance data is increasing exponentially. As more and more baseball insiders begin to rely on sabermetric principles to strategize and evaluate performance, there could soon be a paradigm shift. After all, in only a decade, sabermetrics has gone from relative obscurity to mainstream recognition (Beradino, 2003).

OPPOSITION TO SABERMETRICS

Consistent with any objective evaluations of performance, there are limitations and criticisms of sabermetrics. In fact, the more mainstream sabermetrics become, the more critics and skeptics

will surface. One could not expect to reject over a century of traditionalist mentality about America's Pastime and be received with open arms, neither do sabermetricians. Sabermetrics is by no means an overnight phenomenon. It has been over forty-five years since a mechanical engineer from Baltimore named Earnshaw Cook introduced his vision of how baseball should be analyzed and managed. As the result of a lunchtime conversation with a Yale psychology professor concerning the productiveness of the sacrifice bunt, Cook embarked on a three year journey to present a formal analysis of baseball. His analysis suggested that no one had ever known the true percentages of the game, and if anyone did know them he could manage nearly any team to success. As with many innovations, the first edition of *Percentage Baseball* was met with bitter criticism and controversy. In regards to Cook's book, *The Sporting News* writer James Gallagher wrote "I do not understand how the Baltimore mathematicians reached their controversial conclusions, but in my book any generalizations about baseball have to be wrong" (Schwarz, 2004).

Even Cook himself understood the difficulty of introducing his ideas to a skeptic and unwilling audience. Cook (1966) wrote the following in his forward to *Percentage Baseball*:

> "The general complacency of baseball people-even those of undoubted intelligence-toward mathematical examination of what they regard properly and strictly as their own dish of tea is not too astonishing. I would be willing to go as far as pretending to understand why none of four competent and successful executives of second-division ball clubs were most reluctant to employ probabilistic methods of any description ..... but they did not even want to *hear* about them" (p. xi).

It would take another thirteen years for another book to be published defying conventional baseball logic. In 1977, Bill James self-published a book titled *1977 Baseball Abstract: Featuring 18 Categories of Statistical Information That You Just Can't Find Anywhere Else*. Unlike Cook's Book, James' did not stir up a controversy. This was largely in part because almost no one bought it. The book sold only seventy five copies. James followed up the failure of his first book with another the following year titled *1978 Baseball Abstract: The 2nd Annual Edition of Baseball's Most Informative and Imaginative Review*. It sold only 250 copies (Lewis, 2003).For the next ten years James published annual books in the series of abstracts, each garnering more attention and inviting more criticism. In addition, others began to write and publish their own theories and analytical analyses of the game, some more popular than others (Schwarz, 2004).

Most criticisms of sabermetrics come from baseball insiders, fans, and the media. Critics of sabermetrics suggest that far too much emphasis is placed on mathematical formulas and advanced statistical equations. Traditionalists maintain that it is impossible to compute the human element involved in the game, such as strategy, player management practices, or player engagement. Further, those opposed to the sabermetric movement argue that the counting statistics such as, runs

batted in, home runs, batting average, stolen bases, etc. are exact and incapable of misinterpretation (Albert & Bennett, 2003).

DETERMINANTS OF VALUE

According to Harder (1992), sabermetrics represent an individual's overall contribution to his team, rather than just one element. Sabermetrics supports the notion that a player's value, as it relates to performance, is not dependent on one-dimensional aspects of his game, such as his ability to hit for power, his speed, his propensity for avoiding defensive errors, or his ability to hit for average, nor is it indicative of external factors, such as his spot in the batting order or the performance of his teammates. Rather, according to sabermetric logic, a player's value should be tied to his contribution to his team and his contribution to winning games. Sabermetrics measures this, traditional statistics do not. Given these two dramatically different approaches to evaluating performance, one might expect to encounter disparities when comparing the two measurements to one another.

Research has proven that traditional statistics and sabermetrics can paint vastly different pictures of a player (Bialik, 2003). For instance, when assessing the 2002 offensive performance of Cleveland Indians first baseman, Jim Thome, the traditional measurement of batting average reveals that the four-time all-star hit just .266, 38 points off his 2001 performance, indicating a decline in offensive production. However, by the measure of OPS and Runs Created, Thome put up MVP like numbers, ranking 2nd in all of baseball in both categories behind only Barry Bonds.

Similarly, when assessing the 1996 offensive production of Houston Astros first baseman, Jeff Bagwell, traditional statistics show that the seven time all-star and 1994 National League Most Valuable Player ranked just 31st in batting average, 34th in home run production, and 14th in runs batted in, below average productivity at best for the likely future hall-of-famer. However, according to sabermetric measurements, Bagwell out-performed most of the league by ranking 12th in OPS, 6th in Runs Created, and 1st in Win Shares. Likewise, when evaluating the 1999 offensive performance of New York Yankees shortstop, Derek Jeter, traditional measures show that the World Series Most Valuable Player ranked just 65th in home run production and 50th in runs batted in. Yet, according to sabermetric measures, Jeter finished 4th in Runs Created, and 2nd in Win Shares.

When trying to find cumulative disparities between traditional and sabermetric measurements for the entire decade of the 1990s, one need look no further than five time National League Most Valuable Player, Barry Bonds. According to traditional statistics, Bond's average ranking for the 1990s was 63rd in batting average, 14th in home runs, and 20th in runs batted in. Based on those numbers alone, one might have a hard time recognizing those rankings as belonging to one of the greatest hitters in the history of the game. However, sabermetrics paint a vastly different picture of Bonds's production in the 1990s. For the decade, Bonds's average ranking was 4th in OPS, 2nd in Runs Created, and 2nd in Win Shares.

Even though sabermetrics and traditional statistics have the same purpose, to measure player performance, traditional statistics are limited in their ability to offer a comprehensive assessment of productivity and value. As a result, both are posited to show different interpretations of player performance. Teams and baseball executives have now begun to recognize these differences between traditional statistics and sabermetrics with respect to determining value and measuring productivity. Motivated mostly by economic necessity and the pursuit of competitive advantage, many organizations have embraced change and adopted sabermetric philosophies. As a result, the nature of the sport and the business of professional baseball is rapidly changing.

Of course, some traditional baseball purists will continue to scoff at the notion of a "new and perhaps "improved" way to analyze the game to assess player performance. Most critics of sabermetrics would probably subscribe to the old cliché: "if it ain't broke, don't fix it." However, professional baseball is a business, and a very profitable one at that. Thus, it is understandable why organizations would want to utilize and employ the best strategies available when making personnel decisions that ultimately affect the bottom line: money.

It should be noted that while sabermetrics is not an exact science, it does have its place in the game. The purpose of sabermetrics is not to replace traditional statistics all together. Also, sabermetrics does not attempt to account for the intangible and unpredictable aspect of human behavior. What is does is provide a better evaluation of player performance while offering a more comprehensive look at the intricacies of the game that traditional counting statistics do not measure. In an objective baseball world, one could potentially conceive of a perfect union between traditional baseball wisdom, sabermetric measurement, and old fashioned "gut instinct." For now, one will just have to settle for controversy.

REFERENCES

Albert, J., & Bennett, J. (2003). *Curve Ball.* New York, NY: Copernicus Books.

Alchain, A., & Demsetz, H. (1962). Production, information costs, and economic organization. *American Economic Review, 6,* 777-795.

Associated Press, (2002, November 30). Meet Boston GM Theo Epstein. *USA Today*. Retrieved December 9, 2004, from http://www.content.usatoday.com/sports/baseball/ al/redsox/2002-11-30-epstein-feature_x.htm

Barry, R. R. (1988). *The Application of sabermetrics to the teaching and coaching of collegiate baseball*. Doctoral dissertation, Middle Tennessee State University. (UMI No. 8905090)

Berardino, M. (2003, March 21). The great debate: While Sabermetrics have made great inroads in the game, some still view statistical analysis with skepticism. *Baseball America.* Retrieved September 9, 2004, from http://www.baseballamerica.com

Bialik, C. (2003, July, 1). Sabermetrics goes mainstream. *Wall Street Journal Online* Retrieved September 18, 2004, from http://www.wsj.com

Birger, J. (2003, April 1). Baseball by the numbers: Relying on data made Red Sox owner John Henry a successful trader; Will the same approach work for his team? *Money*. Retrieved April 30, 2004, from http://www.highbeam.com/library/doc0.asp?docid=1G1:98614129&refid=ink_puballmags&skeyword=&teaser=

Bishop, J. (1987). The recognition and reward of employee performance. *Journal of Labor Economics, 5*, S36-S56.

Boswell, T. (1985, April). Player's can't hide from total average. *Inside Sports*, *4*, 26-30.

Cook, E. (1966). *Percentage Baseball*. Cambridge, MA: M. I. T. Press.

deMause, N. (2002). The stat-head revolution: Geeks infiltrate baseball's front offices; Conventional wisdom flees. *The Village Voice*. Retrieved October 2, 2003, from http://villagevoice.com

Gillette, G. (1993). *The great American baseball stat book*. New York, NY: HarperPerennial.

Grabiner, D. (n.d.). *The Sabermetric Manifesto*. Retrieved July 23, 2003, from http://www.baseball1.com/bb-data/grabiner/manifesto.html

Hadley, L., Poitras, M., Ruggiero, J., & Knowles, S. (2000). Performance evaluation of National Football League teams. *Managerial and Decision Economics, 21*, 63-70.

Harder, J. W. (1992). Play for pay: Effects of inequity in a pay-for-performance context. *Administrative Science Quarterly, 37*, 321-335.

Hitzges, N., & Lawson, D. (1994). *Essential Baseball.* New York, NY: Penguin Books.

Huckabay, G. (2003, August 8). 6-4-3: Back to basics. *Baseball Prospectus*. Retrieved August 29, 2004, from http://www.baseball-analysis.com

James, B. (1979). *Baseball abstract*. Lawrence Kansas: Self published.

James, B. (1982). *The Bill James baseball abstract*. New York: Ballantine Books.

James, B. (1984). *The Bill James baseball abstract*. New York: Ballantine Books.

James, B. (1985). *The Bill James baseball abstract*. New York: Ballantine Books.

James, B. (1989). *This time let's not eat the bones: Bill James without the numbers.* New York: Villard.

James, B., & Henzler, J. (2002) *Win shares.* Morton Grove, IL: STATS Publishing.

Krautmann, A. C. (1990). Shirking or stochastic productivity in Major League Baseball? *Southern Economic Journal, 14,* 961-968.

Lehman, D. (1984, April 23). Ballpark figures. *Newsweek*, *24,* 75-76.

Lewis, M. (2003). *Moneyball: The Art of winning an unfair game*. New York, NY: W. W. Norton & Company.

Neyer, R. (2002, April 11). *Bill James is back with "Win Shares."* Retrieved January 13, 2004 from http://espn.go.com/mlb/columns/neyer_rob/42798.html

Pinder, C. C. (1984). *Work motivation.* Glenview, IL: Scott Foresman.

Quinn, T. J. (2003, July 13). Baseball's new magic number. *New York Daily News.* Retrieved October1, 2003, from http://www.nydailynews.com

Schwarz, A. (2004). *The numbers game: Baseball's lifelong fascination with statistics*. New York, NY: St. Martin's Press.

Surowiecki, J. (2002, September 23). The Buffett of baseball. *The New Yorker*. Retrieved May 3, 2004, from http://www.newyorker.com/talk/content/?020923ta_talk_Surowiecki

Surowiecki, J. (2003, June 10). Moneyball redux: *Slate* talks to the man who revolutionized baseball. *Slate.* Retrieved July 17, 2004, from http://slate.msn.com/ id/2084193

Thaler, R. H., & Sunstein, C. R. (2003). Who's on first. *The New Republic*. Retrieved August 20, 2004, from http://www.law.uchicago.edu

Thorn, J., & Palmer, P. (1985). *The hidden game of baseball*. Garden City, NY: Doubleday.

Thorn, J., & Palmer, P. (1990). *Total baseball.* New York, NY: Warner Books.

Woolner, K. (1997, January 18). The First Stathead. Message posted to Boston Red Sox electronic mailing list, archived at http://www.stathead.com/bbeng/woolner /brickey.htm

SPATIAL DIVERSIFICATION: THE CONCEPT AND ITS APPLICATION TO GENERAL GROWTH PROPERTIES INVESTMENT PORTFOLIO

Mark R. Leipnik, Sam Houston State University
Gang Gong, Sam Houston State University

ABSTRACT

The concept of spatial diversification will be introduced and the use of the spatial analysis capabilities of geographic information systems to map, study and derive a series of quantifiable measures of the degree of clustering versus spread of locations associated with various investments will be presented. The benefits and detrimental aspects of having a spatially diversified set of investments will be discussed. Finally, the investment portfolio of General Growth Properties (GGP) the largest real estate investment trust (REIT) to fail in the current economic downturn will be examined and analyzed with respect to the degree to which a lack of spatial diversification may have contributed to its failure. The practical applicability of using any of several measures of spatial diversification as a tool for evaluating the risk associated with specific investment decisions and the issue of which types of investments the concept/method is most applicable will also be recounted.

INTRODUCTION

The concept of spatial diversification is that an investment portfolio whose component investments are located in spatially disparate locations will be more diversified and hence less likely to suffer declines in a down market than an investment portfolio that is more tightly clustered in terms of the locations of its component investments. While in theory this might appear to be a reasonable argument it is important to both test the theory and also determine exactly how and when spatial diversification might be determined. In order to use the spatial analysis and mapping tools of geographic information systems (GIS) to determine spatial diversification it is beneficial to delve into a specific example. This paper will examine the spatial distribution of the large shopping mall developer General Growth Properties

ANALYSIS OF SPATIAL DIVERSIFICATION OF INVESTMENTS

Choosing an investment portfolio upon which to test the concept of spatial diversification is a challenge. One must choose investments where spatial locations can be specifically and accurately determined, where those locations are fixed over a considerable period of time, where there are a reasonable number of locations involved and where the location of an investment is an important factor in determining the return that the investments are likely to yield. Some investment portfolios do not lend themselves to a spatial analysis. For example investments in U.S. Treasury securities cannot be located beyond the entire United States, investments in stock of a multinational corporation like Coca-Cola which does business in millions of locations and has major capital investments in over 180 countries do not lend themselves to analysis using spatial considerations (Coca-Cola, 2010). Also a business that is not tied to a location such as the entertainment industry or oil field services firms are a poor choice for delving into the concept of spatial diversification. A firm or portfolio that involves many disparate types of investments such as the California Public Employees Retirement Fund is also a poor choice since there are just too many investments to map or analyze spatially. An example of an ideal business to analyze spatially is a shopping center development firm and/or investment fund and particularly a shopping mall REIT like GGP. In this type of investment the physical location of the real estate is fixed, the location of the shopping mall can be easily determined and mapped (to within a few meters of the center of the largest building on the mall pad) and the location is very important to the success or failure of the business. The investments are also very large so the duration of the investment is likely to be long and the number o properties involved limited. Since the cost of developing, expanding or acquiring a shopping mall is large even the largest firms in the industry have only a few hundred malls which makes the rather tedious process of geocoding (putting the location of the mall into the GIS as a new layer of data that can then be analyzed) is manageable (Price, 2008). Once the shopping malls have been geocoded and placed into a GIS a variety of spatial analysis tools can be used to analyze the clustering of mall locations, the proximity of malls to each other and the location of malls in relation to the demographics and physical infrastructure at that locality and in the surrounding State, MSA, Zip code zone, census tract or spatially determined trade area.

CASE OF GENERAL GROWTH PROPERTIES

Selection of GGP

Starting with the premise that analysis of firms investing in the development of shopping centers was a good potential basis for developing and testing a model of spatial diversification, the authors began to examine some of the firms involved in this industry, some like Trammell Crow (Trammell Crow, 2010) and Weingarten Realty Investors (Weingarten, 2010) also invested in many

other categories of commercial real estate, others like the United Investors Reality Group (UIRG, 2009) are regional in character (investing only in Arizona, Florida and Texas) and hence in their very investment premise were somewhat clustered (albeit Texas is a physically large state). However, one firm rapidly emerged as an excellent initial case study of the applicability and limitations of the concept of spatial diversification. That firm was General Growth Properties (GGP). This Chicago Illinois headquartered firm founded in 1954 with its first mall development being the Town and Country Center in Cedar Rapids Iowa, besides became the second largest shopping mall developer in the world by 2008. GGP has one outstanding characteristic that led to its selection. It is in bankruptcy. The filing took place on April 17, 2009 (GGP, 2009). At that time GGP stock was trading at .48 cents a share and was delisted from the New York Stock Exchange. GGP stock reached an all time high of 44.23 per share on March 16, 2008. In March 2010 it was relisted on the NYSE and is now trading around $15 per share down from a high in 2009. Although beleaguered Simon Group a rival REIT is considering acquiring the ailing firm (Hudson 2010, April).

Figure 1. GGP Mall Locations.

If the idea is to test the concept of spatial diversification as a means of reducing investment risk, than examination of the spatial distribution of the investments of a firm that has failed by the fairly objective criteria of having to seek bankruptcy protection from creditors bent on the firms

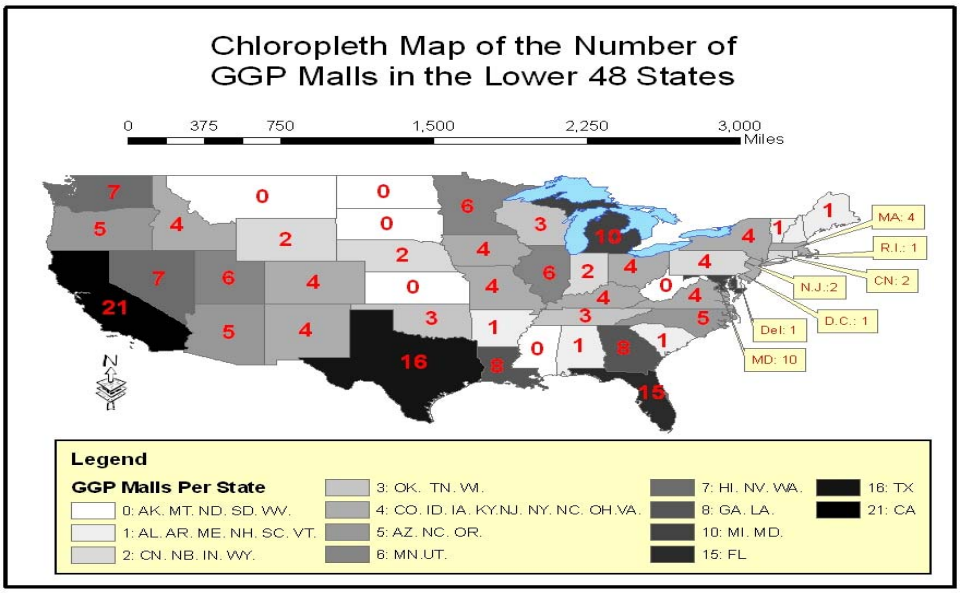
liquidation and dismemberment, seems to be as good choice. It might even be a better choice than study of a highly profitable firm, since those profits might be fleeting or illusory as in the case of Enron (Fox, 2003). This is particularly true in the case of GGP since it is the largest REIT to fail in the current economic downturn. Ironically, for a financial crisis and recession that has generally been attributed by economists to problems in real estate markets, relatively few REIT's have failed. But GGP is a big exception. The authors therefore propose to map the distribution of GGP investment properties and attempt on the basis of their locations to if not to prove that a lack of spatial diversification contributed to the decline of the firm, then to use this examination as a start to longer the process of fleshing out the concept of spatial diversification in order to examine spatial factors and limitations of the concept using the test case of a real real-estate firm that was open to investors and therefore has extensive annual reports and financial statements that are readily available to researchers.

GGP's Investment Strategy

Growth is (or rather was) GGP's mantra and middle name. The firm founded in 1954 aggressively grew in the 1990's and 2000's by acquisitions and new green field projects during the particularly during the mid to late 2000's. The firm concentrated on investments in figuratively and often literally hot markets in the Sun Belt such as California Nevada, Florida, Hawaii, Arizona, and Louisiana. The firm also invested heavily in the rust belt areas of Michigan and in Maryland, In addition, the firm invested in many properties in Texas and California. These two states have ~10% and ~7% of total U.S. population and represent ~10% and 7% respectively of all GGP malls. Therefore on this simple basis one might conclude that GGP did not disproportionally invest in either Texas or California. When one adjusts for population of each state, the states that stand out in particular as over represented are Hawaii with 7 malls including the billion dollar (book value) Ala Moana Center (GGP's most valuable property), Nevada also seems over represented with 7 malls (all located in Las Vegas), the 15 malls in Florida, the 10 malls in Michigan and the 10 malls in Maryland are all disproportionate to population of each state. Some small population states such as Idaho (with 4 malls) and Wyoming (lowest population state in the U.S.) with 2 malls are less spectacular examples of disproportionate investments. Conversely, there are only 4 malls in each of Ohio, Pennsylvania and New York while in contrast the much smaller population state of Utah has 6 GGP malls. This is harder to understand than the lack of properties in Alaska, Montana, West Virginia, and North and South Dakota. The lack of properties in Kansas is a surprise. Figure 1 shows the location of all of GGP's mall properties in the United States, while Figure 2 shows the number of malls per state with a chloropleth color fill indicating relative populations of each state in the lower 48 states (it omits the 7 malls that GGP owns in Hawaii). In addition, to investing in dispersed locations throughout these states, GGP choose to place a very disproportionate amount of its

investments in relatively few prestige properties in places like Las Vegas, Orlando, Florida, Towson, Maryland, Austin, Texas and in Hawaii.

Figure 2. GGP Shopping Malls per State.



Spatial Analysis of GGP Properties.

Based on population per mall for the entire U.S. not including Alaska, there should be one mall for every 1.4 million persons. Figure 3. shows the difference between the number of actual malls in 2008 and the number that would be predicted based on 2000 population estimates of each state. One form of spatial analysis is to determine the geographic center of a scattered set of points for all 219 malls. This is shown in Figure 5. And the results of the spatial analysis are presented in Table 1. Which lists the results from the Average Nearest Neighbor Distance analysis on GGP's 219 shopping malls. The results show that the nearest neighbor ratio for the 219 GGP's shopping malls is 0.41 (less than one indicating clustering pattern while greater than one suggesting dispersion) with a p-value of 0.0000 (highly significant). This indicates that there is a distinct clustering pattern among GGP's shopping malls nationwide. Of the 219 malls currently owned by GGP, 20 properties have been designated platinum properties, these are shown in Figure 6. These are the malls which GGP has invested the most money into. To facilitate further spatial analysis of GGP's investment, a GIS layer showing the locations of GGP's platinum shopping malls is necessary. This was done first by extracting the street addresses of all the shopping malls from GGP's website, then their

longitude-latitude coordinates were obtained by converting street addresses through an online geocoding service, and last the coordinates were fed into ESRI's ArcGIS to generate a point layer in which each shopping mall is represented by a point (or a star as shown in Figure 6). A quick glance of the distribution of these flagship properties reveals that they are even more clustered than the overall pattern of GGP malls. Of the 20 premium properties, 3 are in Hawaii and 4 in Las Vegas. However, a convincing conclusion regarding clustering vs. dispersion requires quantifiable measurements.

Figure 3. Actual and expected number (based on population of each State) of GGP malls.

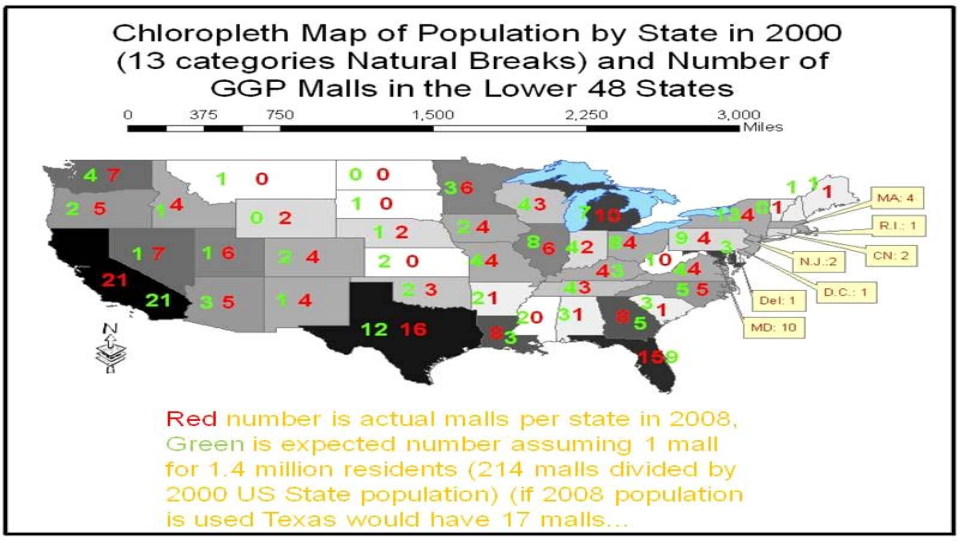


Figure 4. Mean Geographic Center of all 219 GGP malls.

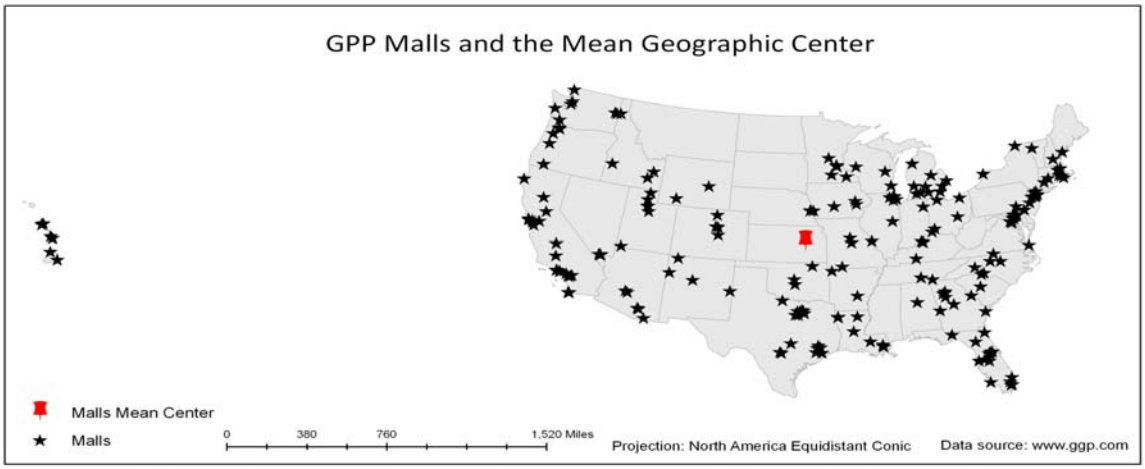
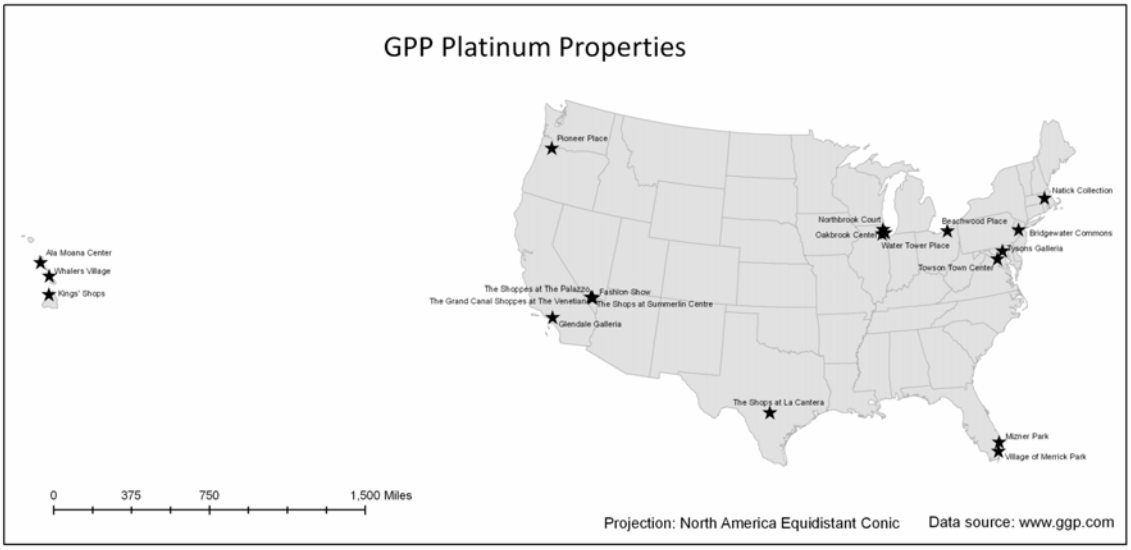


Table 1: Average Nearest Neighbor Analysis for All GGP Malls in the USA in 2008	
Observed Mean Distance	56648.400046
Expected Mean Distance	137166.718493
Nearest Neighbor Ratio	0.412989
Z Score	-16.656670
p-value	0.000000

Figure 5. GPP Platinum Properties.



METHODS FOR ANALYSIS OF SPATIAL DIVERSIFICATION

There are a variety of methods that can be used to examine the clustering of dispersion of spatial locations. These include density (also called heat) mapping which can determine "hot spots" that are locations with an above average concentration of some phenomena (in this case shopping malls (Bachi, 1993). Adjusting both for the spatial proximity or clustering of malls and the clustering of population that might patronize malls will help to better understand areas where there is an unjustified concentration of malls. When this is done for a state, Hawaii, Nevada, Michigan, Utah and Idaho stand out as states with a high density of malls but not a concentration of population. One must then examine the clustering at a more detailed spatial scale. In Nevada all the malls are in Las Vegas, but so is 80% of Nevada's population. In any case the resident population is not the only source of shoppers. In a place like Las Vegas and Hawaii and in areas like Orlando Florida one should figure in the tourism (and in Nevada gaming) related traffic. In fact the drop in vacation travelers patronizing GGP's 5 new and very pricy and recently developed malls in Nevada (in

addition to the long standing Meadows and Boulevard malls) is undoubtedly a major factor in the firm's financial plight.

In addition to a hot spot also called a density analysis one can use a GIS to measure distances between features like mall locations represented as points and stored as a layer of data in the GIS. One can measure the distance between each pair of malls (allowing for the curvature of the earth using an equidistant projection). When summed for all the pairs of malls a larger total distance would indicate a greater degree of separation and in the simplest case a greater degree of spatial diversification. The actual measure of distance can be a Euclidean (straight line) distance which is easiest to calculate. The distance can also be based on the shortest path over a road network (a so called Manhattan distance) (ESRI, 2007). The actual travel time can be used as well although some assumptions about mode of travel and speed limits need to be made in this case. The GGP case is illustrative of one issue with this approach which is that for some malls the mode of travel will differ. Normally this would not be too problematic but in GGP's special case many of the most costly investments were in malls not generally accessible by motor vehicles but rather by persons who relied on air travel to reach the city in which the Mall was located from their homes. This is certainly true for the malls in Hawaii, also the ones in Las Vegas and Orlando. Therefore the best, but hardest to determine measure of distance between malls would be the time or even the cost that typical customers incurred in travel to and from the mall. Just determining the spatial distance between GGP's properties would not determine if the spatial distribution was clustered or dispersed, but it would allow analysis of the degree of clustering of GGP's properties versus that of other firms. Another distance-based measure of spatial diversification explored by the authors can be called the total spatial deviation (from the centroid). Similar to the statistics of variance in which the data dispersion is measured by the deviation from the mean, a spatial deviation of a shopping mall can be defined as its distance from the geographical centroid of all the properties. If all the shopping malls are spatially dispersed, their total spatial deviation would be larger than that of a group of clustered shopping centers.

The simplest method to find the geographical centroid is through calculating the average coordinates of all the 20 platinum properties. ArcGIS Spatial Statistics Tools offer a similar algorithm to calculate the "mean center" which identifies the geographic center of a set of features. Figure 6. shows the location (indicated by the sign of a pushpin) of the centroid of GPP's platinum properties. The weighted centroid can also be obtained by applying a certain measure of weight on each property. For example, Figure 7. shows a proportional symbol map of GPP's platinum properties with the size of the symbol proportional to the property's total leasable area. If the total leasable area at each shopping mall is factored in as weight, Figure 8. illustrates the new weighted center – dragged a little to the east by the relatively bigger shopping malls in Chicago. Once the geographical centroid is located, the total spatial deviation of all the platinum properties can be easily calculated to assess spatial diversification. While being easy to calculate and use, total spatial

deviation has its own limitation. Its lack of robustness becomes salient when it fails to differentiate the following two scenarios shown in Figure 9.

Figure 6. The mean center (centroid) of GPP's Platinum Properties.

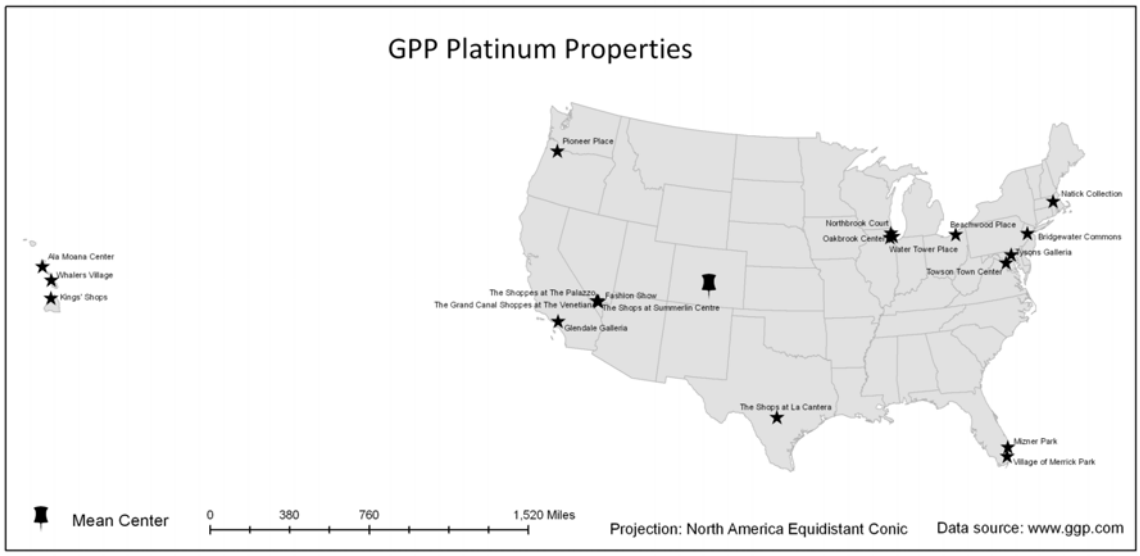


Figure 7. Total Leasable Area at GPP's Platinum Properties. (Symbol size is proportional to the total leasable area in square feet)

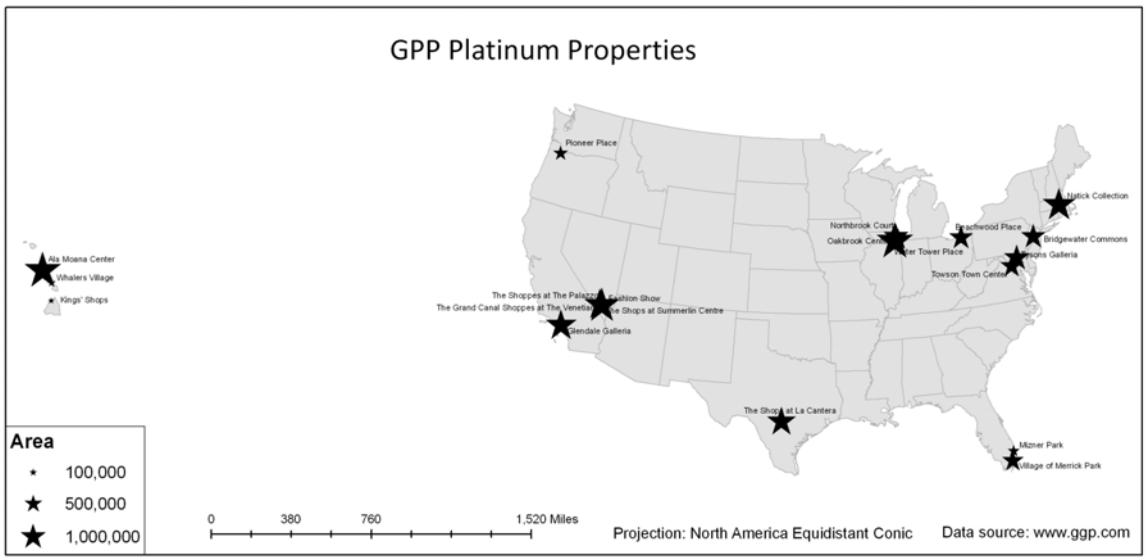


Figure 8. Weighted mean center (centroid) by total leasable area of GPP's Platinum Properties.

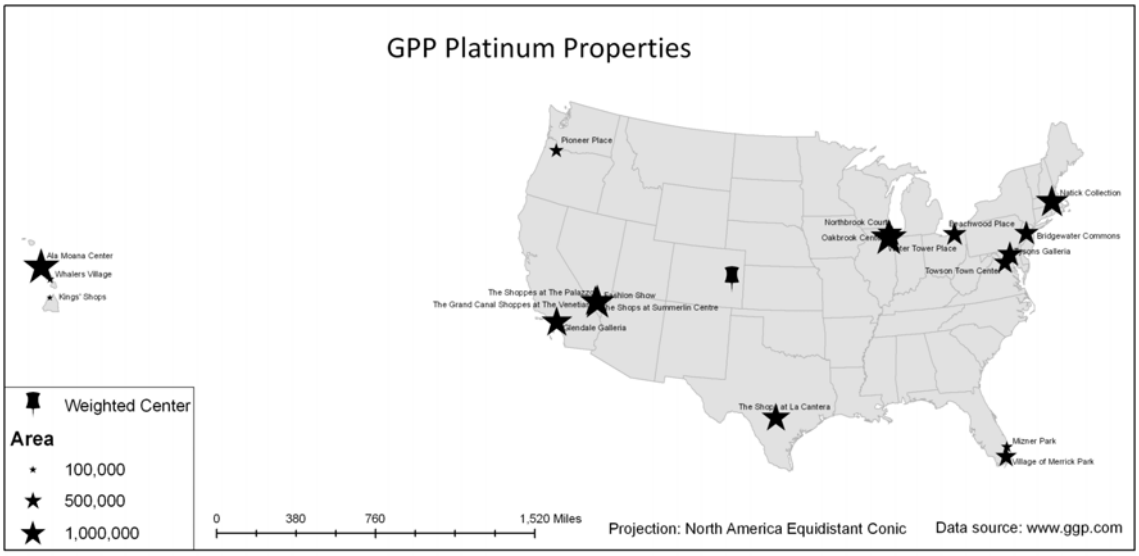
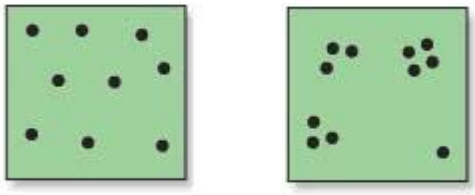


Figure 9. Two distributions with different patterns, yet their spatial deviation are very similar.



Average Nearest Neighbor Distance

To overcome the problem illustrated in Figure 9, another distance-base measure of spatial diversification named Average Nearest Neighbor Distance was investigated by the authors. The Average Nearest Neighbor Distance measures the distance between each feature and its nearest neighbor's location, and then averages all these nearest neighbor distances. Further this actual average nearest neighbor distance is compared to a hypothetical average nearest neighbor distance under the assumption of random distribution. If the actual average distance is smaller, the distribution of the features being analyzed are considered clustered. If the average distance is greater than a hypothetical random distribution, the features are considered dispersed. The index is

expressed as the ratio of the observed distance divided by the expected distance (expected distance is based on a hypothetical random distribution with the same number of features covering the same total area). Table 2 lists the results from the Average Nearest Neighbor Distance analysis on GGP's platinum properties. The results show that the nearest neighbor ratio for GGP's platinum properties is 0.56 (less than one) with a p-value of 0.000195 (highly significant). This indicates that there is a clear clustering pattern among GGP's platinum shopping malls.

Table 2.: Average Nearest Neighbor Analysis for Ggp Platinum Properties	
Observed Mean Distance	245550.279509
Expected Mean Distance	434902.681458
Nearest Neighbor Ratio	0.564610
Z Score	-3.724986
p-value	0.000195

Important Caveats

Spatial analysis of an investment portfolio like GGP's requires several adjustments to the most simple spatially based analysis. An important caveat to the analysis of travel distance between malls is that one or two outliers (or even 7 malls in Hawaii) may not make a set of investments spatially diversified. Adding a single mall in a far flung point of the compass (for GGP that would be opening a new mall in North Dakota for example) would add a large number of long travel lines between the new mall and existing properties, but that would not really impact the existing tight clustering in places like Maryland (10 malls in a small State). The distance analysis would cross state lines so while there may not be a lot of malls in any one State in New England but the States are small and the one mall in Maine the only larger state is located in southern Maine close to 8 other malls in this small region.

A more critical consideration is that not all investments are of equal magnitude and hence equal importance to the financial risk involved. GGP owns a shopping mall in Rocksprings, Wyoming (White Mountain Mall) that is more a glorified strip center (it has a small super market, a mid-sized sporting goods store and a few other shops and businesses in the pad). It also owns the Boulevard mall in Las Vegas and the Ala Moana center in Hawaii and the Towson Town Center in Maryland and Water Tower Place in Chicago, each mall is the most important in the State. . In fact GGP owns the seven largest malls in the Hawaiian Islands. If one gives equal importance to purchase of a shopping mall in Rocksprings, Wyoming, an investment of less than 10 million dollars to development of the shopping mall GGP values at over $1 billion in Honolulu, Hawaii, then the concept of spatial diversification cannot be tested. One must take into account that each one of the 5 malls GGP invested in Las Vegas cost many times more than the average or even the most expense

mall in most of the other states GGP is involved in. Thus one should weight the analysis of diversification not only by some measure of distance apart but also by the magnitude of the investment. When this is done the already unbalanced investment strategy of GGP begins to show signs of the calamity that befell it. The investments in the Hawaiian Islands also illustrate a limitation of a too simple minded application of distance as a factor in spatial diversification. The GGP malls in Hawaii are on Oahu, Maui and Hawaii, these islands are spaced hundreds of miles apart. But effectively one can ignore the ocean separating them in the analysis because in a sense their proximity is greater than the distance separating them might appear. With respect to factors involving the success of a shopping mall each is largely dependent on a similar set of tourists and local residents and in fact perhaps on the exact same tourist that visits more than one island.

Beyond Distance

While spreading investments into far separated markets might seem to reduce risk, the case of GGP also shows another important limitation on this simple assumption. GGP invested in resort area shopping malls in South Florida, and Orlando, in heavily tourism dependent areas like Las Vegas and in Hawaii. At first blush this might indicate spatial diversification. Hawaii is about as far from Florida as it is possible to get in the U.S. and Nevada is not really near any area except Southern California. The fallacy in this argument is that each area shares many characteristics in common that is dependence on more or less conspicuous and certainly discretionary consumption by resort goers and persons with vacation, time-share, retirement and second homes in these areas. These consumers make purchases from prestige shops in a high end mall that is dependent on their having discretionary income. The economic decline hurt these areas more than most areas in the United States. GGPs strategic focus was on the areas hardest hit by the downturn: GGP concentrated the bulk of its investments into 7 markets, California, Nevada, Hawaii, Maryland, Michigan, Florida and Texas. How does this distribution of investments correlate with the areas that have fared well or poorly in the current economic downturn?

Five out of Seven Can Be Wrong

Amazingly 5 of the 7 states which GGP concentrated is investments are the ones hardest hit by the recession. These roughly in order of severity of decline are Nevada, Michigan, Hawaii, California and Florida. Only Texas and Maryland are not among the states most affected by the recession. GGPs recent large investments were made it the city with the largest drop in real estate values and highest foreclosure rate: Las Vegas, Nevada. One of the better yardsticks of the downturn with regard to real estate related investments is foreclosure activity. GGP's single largest recent investments were in the city of Las Vegas, the most economically distressed real-estate market in the nation. GGP also made multiple investments in states like Maryland where it did develop the

Towson Town Center, a shopping mall in an area that is also doing better than average and some of the smaller malls (such as the mini-mall) in Wyoming is doing fairly well. This is not a coincidence but it is not the outcome of a conscious desire to invest in what would soon become hard hit markets. The 5 out of 7 match between GGP's major investments and failing real estate markets at the state level basically stems from the same source.

The Search for Fast Growth

GGP wanted to be in the hottest markets around the country, ones driven by rapidly raising real estate values, lots of conspicuous consumption financed by home equity lines of credit and flipping real-estate and ones where luxury vacations and resorts were the norm. This matched their focus on high end retailers, rapid growth in revenues and other financial and image related factors. It was fun to invest in Vegas and Hawaii and gullible investors were impressed by glitzy properties and rapid growth in value of buildings and land in hyper inflated local real-estate markets. Notable slow growing but sustainable markets such as Nebraska, Kansas, and other areas like the Mid Atlantic States were neglected because there was neither population growth nor exuberant tourism. Even in the Middle West the trend was followed with a mall in Branson, Missouri. The investment concentration in Michigan is a little harder to explain but it is clear that GGP did not anticipate a catastrophic downturn in the auto industry and related manufacturing that hit this state hard. In any case GGP did not invest nearly as much into these slower growing markets as it did into Nevada, Hawaii, Florida, and California.

CONCLUSION

GGP's investment strategy was a colossal failure; the rapid growth it espoused was dependent on investment in risky markets. While these markets were spatially separated within the selected markets investment was clearly clustered and GGP was the dominant player in Hawaii and Nevada and Utah. These states along with Florida and California had a long and spectacular boom; they also perhaps predictably suffered an even sharper bust. Analysis of the spatial distribution of GGP's investments indicated that it was not diversified, not diversified in terms of close proximity of multiple high value malls in Las Vegas and Hawaii and not diversified in terms of investments throughout the U.S. including in slower growing and less exotic or resort oriented areas. The effort to conduct a spatially based analysis of the GGP investment portfolio indicates that this method has promise, but it requires careful application and adjustment. Specifically, the analysis of locational clustering and distance needs to be weighted with the value of the investment. Also one needs to consider other factors besides distance and proximity since investments in far flung resort properties or investments in far flung automotive plants or any investment that is concentrated in a particular industry or depended on the spending habits of a particular narrow class of consumer no matter how

far separated will be riskier than investments that are balanced. One needs to balance investments both by spatial separation and by differentiation into various industries in order to achieve true diversification.

REFERENCES

Bachi, R. (1993). Standard Distance Measures and Related Methods for Spatial Analysis *Papers in Regional Science. 23(1)* December 1993

Coca-Cola (2010). Company History & Operations. Retrieved April 10, 2010 from: http://www.thecoca-colacompany.com/

ESRI (2007 Calculating Manhattan Distances with ArcGIS. Retrieved April 14, 2010 from: http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Calculate_Distance_Band_from_Neighbor_Count_(Spatial_Statistics)

Fox, L (2003 *Enron: the Rise and Fall*, Chapter 5, pp 78-97 John Wiley.

GGP (2010). *GGP Company History.* Retrieved March 7, 2010 from: http://www.ggp.com/

Price, M (2009) Chapter 10 on Geocoding in *Mastering ArcGIS, 4th ed.* Prentice Hall.

Real Estate Channel (2008). *Rise in Stock Price Puzzles GGP Officials Retrieved.* March 23, 2010 from: http://www.realestatechannel.com/us-markets/commercial-real-estate-1/general-. growth-properties-general-growth-properties-stock-price-the-rouse-new-york-stock-exchange-general-growth-properties-debt-load-finkelstein-632.php

Trammell Crow (2010). *Commercial Real-estate locations*. http://www.trammellcrow.com/

United Investors Reality Trust (2010). *Locations of properties* Retrieved March 24, 2010 from: http://www.uirt.com/

Hudson K. (2010, March) Wall Street Journal, March 17, 2010. *Simon Mulls Acquisition of GGP.* by Kris Hudson, Commercial Real-Estate, editor.

Weingarten (2010). *Distribution and type of investment properties*. Retrieved March 24, 2010 from: http://www.weingarten.com/home/

ENTERPRISE RISK MANAGEMENT (ERM) – FAILURE IS NOT AN OPTION

Robert B. Matthews, Sam Houston State University
Ronald J. Daigle, Sam Houston State University
Paul Vanek, Sam Houston State University

ABSTRACT

The Enterprise Risk Management (ERM) framework provides a useful framework for planning, conducting, and evaluating risk management evolutions in a wide range of enterprises. This paper addresses an overview of ERM, including background, conceptual framework, implementation guidance, and thoughts for future consideration.

INTRODUCTION

> *"You want a valve that doesn't leak and you try everything possible to develop one. But the real world provides you with a leaky valve. You have to determine how much leaking you can tolerate."*
>
> *Obituary of Arthur Randolph, January 3, 1996*

In September 1992 (amended 1994) the Committee of Sponsoring Organizations (COSO) of the Treadway Commission published *Internal Control – Integrated Framework (COSO-IC)*, the result of a project begun in 1987 to develop integrated guidance on internal control. This publication presented a common definition of internal control and a framework for evaluating and improving internal control systems. COSO-IC defined internal control as "a process, effected by an entity's board of directors, management and other personnel, designed to provide reasonable assurance regarding the achievement of objectives in the following categories:

- Effectiveness and efficiency of operations.
- Reliability of financial reporting.
- Compliance with applicable laws and regulations (COSO, 1992)."

In addition to the above three internal control objectives, COSO-IC identified five components of internal control (COSO, 1992):

- Control Environment
- Risk Assessment
- Control Activities
- Information and Communication
- Monitoring

The COSO-IC framework gained widespread acceptance. It became the predominant standard used by U.S. companies use to evaluate their compliance with the Foreign Corrupt Practices Act of 1977 (FCPA). According to a poll by CFO Magazine released in 2006 (Shaw, 2006), 82% of respondents claimed they used COSO-IC for their internal control framework. Other frameworks identified by respondents included COBIT (Control Objectives for Information and Related Technology) 33%, AS2 (Auditing Standard No. 2, PCAOB) 28%, and SAS 55/78 (AICPA) 13%.

Following the turn of the millennium, several high-profile business scandals and failures (Enron, Tyco, Adelphia, Peregrine, and WorldCom) led to enactment of the Sarbanes-Oxley Act of 2002 (SOX), which extends the long-standing requirement for public companies to maintain systems of internal control and requires management to certify and the independent auditor to attest to the effectiveness of those systems. COSO-IC became the broadly accepted standard for satisfying those reporting requirements.

In response to accompanying calls for enhanced corporate governance and risk management, in 2004 COSO published *Enterprise Risk Management - Integrated Framework (COSO-ERM),* which defines enterprise risk management (ERM) as a "process, effected by an entity's board of directors, management and other personnel, applied in strategy setting and across the enterprise, designed to identify potential events that may affect the entity, and manage risk to be within its risk appetite, to provide reasonable assurance regarding the achievement of entity objectives (COSO, 2004, p. 2)." COSO-ERM expanded the earlier definition of internal control to provide a more robust and extensive focus on the broader subject of ERM. COSO-ERM expanded the objectives identified in COSO-IC, to include Strategic in addition to COSO-IC's Operations, Reporting, and Compliance (COSO, 2004, p. 3).

COSO-ERM also modified the internal control components identified in COSO-IC, and increased the number from five to eight, as follows (COSO, 2004, pp. 3-4):

- Changed Control Environment to Internal Environment
- Added Objective Setting, Event Identification, and Risk Response
- Retained Risk Assessment, Control Activities, Information and Communication, and Monitoring

While COSO-IC focused on component units within the enterprise, COSO-ERM focuses on the enterprise level and intermediate division or subsidiary levels as well as the individual component units.

The changes in emphasis resulting from these differences between COSO-IC and COSO-ERM are summarized as follows:

COSO-IC	COSO-ERM
Rules-based, bottom-up approach, at least initially.	Top-down, holistic, principles-based approach.
Focus on controls over transactions.	Focuses on risks associated with events.
When used for SOX compliance purposes, does not specifically address operational, strategic or compliance risks not related to financial reporting.	Specifically addresses operational, strategic, and compliance risks as well as financial reporting risks.

Like the earlier COSO-IC framework, the COSO ERM framework is also gaining increasing acceptance as a standard for risk management in various enterprises.

A somewhat different approach was taken by the Casualty Actuarial Society (CAS) in 2003. CAS defined ERM as the "discipline by which an organization in any industry assesses, controls, exploits, finances, and monitors risks from all sources for the purpose of increasing the organization's short- and long-term value to its stakeholders (ERM Committee, 2003, p. 8)." CAS conceptualized ERM as proceeding across the two dimensions of risk type and risk management processes (ERM Committee, 2003, p. 8). The risk types and examples include (ERM Committee, 2003, pp. 9-10):

- Hazard risk (tort liability, property damage, natural catastrophe)
- Financial risk (pricing risk, asset risk, currency risk, liquidity risk)
- Operational risk (customer satisfaction, product failure, integrity, reputational risk)
- Strategic risks (competition, social trends, capital availability)

The CAS risk management process involves (ERM Committee, 2003, pp. 11-13):

- Establishing Context: This includes an understanding of the current conditions in which the organization operates on an internal, external and risk management context.
- Identifying Risks: This includes the documentation of the material threats to the organization's achievement of its objectives and the representation of areas to the organization may exploit for competitive advantage.
- Analyzing/Quantifying Risks: This includes the calibration and, if possible, creation of probability distributions of outcomes for each material risk.

- Integrating Risks: This includes the aggregation of all risk distributions, reflecting correlations and portfolio effects, and the formulation of the results in terms of impact on the organization's key performance metrics.
- Assessing/Prioritizing Risks: This includes the determination of the contribution of each risk to the aggregate risk profile, and appropriate prioritization.
- Treating/Exploiting Risks: This includes the development of strategies for controlling and exploiting the various risks.
- Monitoring and Reviewing: This includes the continual measurement and monitoring of the risk environment and the performance of the risk management strategies.

Other risk frameworks in use throughout the world include (Schanfield & Helming, 2008):

- AIRMIC – Association of Insurance and Risk Managers
- ALARM – The National Forum for Risk Management in the Public Sector (UK)
- AS/NZ 4360:2004 (Australia/New Zealand)
- British Standard 31100
- CoCo – Criteria of Control (Canada)
- Combined Code on Corporate Governance (UK)
- FERMA – Federation of European Risk Management Associations
- Internal Control (Hong Kong)
- IRM – Institute of Risk Management
- ISO 31000 (International Organization for Standardization)
- King Report on Corporate Governance (King 1)
- King Report on Corporate Governance in South Africa (King 2)
- Risk and Insurance Management Society (RIMS) Risk Maturity Model

Risk management expert Felix Kloman defines risks as, "a measure of the probable likelihood, consequences (favorable and unfavorable), and timing of a future event or situation that would affect the company (Kloman, Felix, quoted in Schanfield & Helming, 2008)." Such a definition focuses upon both the downside risk and the upside opportunity.

BACKGROUND

A review of how things have changed since the 1970s provides some perspective as to the significance of risk management:

1970s	End of Vietnam War Yom Kippur War and first Arab oil embargo, 1973 Dow-Jones Industrial Average (DJIA) high of 1011, 1976 Foreign Corrupt Practices Acts (FCPA), 1977 Fall of the Shah of Iran, US Embassy hostage situation, Iranian oil embargo, 1979 Oil increased from $5/bbl to $15/bbl over the decade
1980s	The "Reagan Years" IBM PC, 1981 DJIA low of 776, 1982 Oil $20/bbl, mid-80s DJIA high of 2722, 1987 Stock market crash, 1987 COSO begins research into fraudulent financial reporting, 1987 Fall of Berlin Wall, 1989
1990s	Desert Storm, 1991 COSO-IC released, 1992 Development of the Internet Fall of Barings Bank, 1997 Oil $10/bbl, 1997 COSO concludes research, 1997 Fall of Long Term Capital Management, 1998 First DJIA close over 10000, 1999 Y2K efforts, 1999-2000
2000	DJIA high of 11723 Dot-Com bubble burst
2001	Terrorist attack of 9/11 Fall of Enron Basel II Accords introduced
2002-03	Fall of Arthur Andersen, WorldCom, and Adelphia Sarbanes-Oxley Act of 2002 DJIA low of 7286, 2002 Continued political unrest Global "War on Terrorism" CAS-ERM issued, 2003 Oil $30/bbl
2004	Auditing Standard 2 (AS2) released by PCAOB First year for SOX 404 compliance for large public companies COSO-ERM released Oil $50/bbl
2005-06	AS2 required for external auditors Oil $79/bbl, 2006

2007	Audit Standard 5 (AS5) supersedes AS2 DJIA high of 14164 Oil $99/bbl
2008	Global recession Failure of US financial institutions and TARP response Oil $120/bbl
2009	DJIA low of 7062 Economic stimulus plan Increase in pirate activity in Indian Ocean Christmas Day aircraft bomb attempt over Detroit, Michigan Oil $50-$70/bbl
2010	Blowout of Mississippi Canyon 252 oil well operated by BP Times Square truck bomb attempt

Flowing through those events are the following general trends that must be kept in mind in comparing the need for a more rigorous ERM today than previously:

YESTERDAY	TODAY
Simpler times	Requirements, systems, and tools are more complex
Frequent breakdowns occurred within companies, but repairs could be made without computer scientists, engineers, attorneys, environmental experts, accountants, and financial analysts. Failure in one area of the business seldom directly impacted another area	Breakdowns can lead to a significant "domino effect" with far-reaching consequences
Hazards which ultimately resulted in losses were easier to contain.	Media's role has changed from observer to a catalyst of negative public opinion

In November 2007, Standard & Poor's (S&P) announced a Request for Comment: Enterprise Risk Management Analysis for Credit Ratings of Non-financial Companies (Dreyer & Ingram, 2007). When the comment period closed in March 2008, over 90 responses had been received. The comments generally supported S&P's proposal to introduce ERM analysis for non-financial companies. In May 2008 S&P announced that they would want to include ERM in its evaluation of non-financial companies. During the third and fourth quarters of 2008, S&P worked to develop benchmark and evaluation criteria. In 2009, S&P began to include ERM in its evaluation of credit ratings. S&P focused on the risk management culture and strategic risk management. S&P views ERM as a gauge of the quality of management at the helm.

Also in November 2007, S&P reported on its ERM evaluation process for insurers (Santori, Bevan, & Myers, 2007). This reflected the results of a pilot program conducted by S&P that included 78 insurance companies. The composition of the pilot companies was 37% property and

casualty, 21% life, 13% reinsurance, 12% health, 12% multiline, and 1% mortgage insurers. The S&P ratings breakdown was 13% AA and AAA, 45% A, and 42% BBB and lower. S&P found the quality of risk management to be as follows (Santori, Bevan, & Myers, 2007):

8%	Excellent	Master of controls, preparations for unknown future risks, and strategic applications
24%	Strong	Basic risk controls in place for all major risks, plus processes to prepare for unknown future risks and to make strategic choices among risks based on risk/reward framework
62%	Adequate	Basic risk controls in place for all major risks
6%	Weak	Lacking basic controls for important risk(s)

S&P also found that in assessing the ERM impact on ratings (Santori, Bevan, & Myers, 2007):

5%	ERM evaluations strengthened the ratings
25%	ERM evaluations affirmed or supported ratings
65%	ERM evaluations were neutral to ratings
5%	ERM evaluations were negative to ratings

Comparison of these two sets of results suggests a possible correlation between findings and ratings, as follows:

Findings		Impact	
Excellent	8%	5%	Strengthen
Strong	24%	25%	Affirm/support
Adequate	62%	65%	Neutral
Weak	6%	5%	Negative

The crisis in the global banking industry provides an obvious recent example of the consequences of failure to assess enterprise risks effectively. This sector once claimed leadership in risk management. That reputation has been lost in a flurry of bad loan portfolios, failed banks, nationalization/bailouts of some banks, and shotgun mergers of others. Several prominent organizations have weighed in with analyses of what went wrong (Baker, 2008).

The Financial Stability Forum (FSF) issued a report on 2 April 2009, "FSF Principles for Sound Compensation Practices," which stated in part (FSF, 2009):

- "Compensation practices at large financial institutions are one factor among many that contributed to the financial crisis that began in 2007. High short-term profits led to generous bonus payments to employees without regard to the longer-term risks they imposed on their firms. These perverse incentives amplified the excessive risk-taking that severely threatened the global financial system and left firms with fewer resources to absorb losses as risks materialized. The lack of attention to risk also contributed to the large, income cases extreme, absolute level of compensation in the industry."
- "To date, most governing bodies (henceforth, 'board of directors') of financial firms have viewed compensation systems as being largely unrelated to risk management and risk governance. This must change."
- "As a practical matter, most financial institutions have viewed compensation systems as being unrelated to risk management and risk governance."

Shortly after the FSF report, the Institute of International Finance (IIF) issued a report stating that the crisis "raised questions about the ability of certain bank boards to oversee senior managements and to understand and monitor the business (Baker, 2008)."

The Economist Intelligence Unit (EIU) surveyed banks worldwide and reported that only 18% had an ERM strategy in place that was "well-formulated and rolled out across the business (Baker, 2008)."

The Association of Chartered Certified Accountants (ACCA) reported that (ACCA 2008):

- The principal source of the global credit crunch is a failure of corporate governance at banks, which encouraged excessive short-term thinking and blindness to risk.
- Risk management and remuneration/incentive systems must be linked. Executive bonus payments should be deferred until there is incontrovertible evidence that profits have been realized, cash received, and accounting transactions cannot be reversed.

Bruce Caplain has identified three factors that are imperative in an enterprise's ERM effort (Caplain, 2008):

- Management's commitment, including the Board
- The enterprise's governance structure of oversight functions that focus on risk and on identifying and mitigating issues
- The design of the enterprise's ERM effort—is it just another program, or is the risk mind-set fully embedded?

As suggested by the foregoing, the value of ERM may be at its greatest during times of economic decline or crisis. Several factors operate:

- The changing risk environment (KPMG 2008)
  - Arguably there have never been more risks to a business than there are in the current marketplace.
  - Even leaving aside today's prevailing concerns around the credit crunch, consider the following:
    - Technology entering new markets
    - Changing consumer habits
    - New products
    - Dealing with emerging economies.
  - These are all aspects of business which carry far greater risks than they used to, thanks to the effects of globalization and a more demanding end-user.

- Increased scrutiny from legal and regulatory agencies (Wheeler & Yoo 2009)
  - SEC
  - Department of Justice
  - Stock exchanges
  - Securities fraud trial lawyers
  - Sections 302 and 404 of the Sarbanes-Oxley Act
  - Foreign Corrupt Practices Act of 1977
  - Industry-specific regulations (privacy, anti-money-laundering, risk-based capital requirements)

- Increased criticism from shareholders and other stakeholders (Wheeler & Yoo 2009)
  - Outsourcing/third party resources
  - Credit rating agencies
  - Institutional investors
  - Personal liability for Board members

CONCEPTUAL FRAMEWORK

The COSO-IC, COSO-ERM, and CAS-ERM structures share many common elements. The CAS-ERM risks can be related to the COSO-IC and COSO-ERM objectives, as follows:

COSO-IC OBJECTIVES	COSO-ERM OBJECTIVES	CAS-ERM RISKS
	Strategic	Strategic risk
Operations	Operations	Operational risk
Financial Reporting	Financial Reporting	Financial risk
Compliance	Compliance	Hazard risk

Similarly, the CAS-ERM process can be related to the COSO-IC and COSO-ERM components, as follows:

COSO-IC COMPONENTS	COSO-ERM COMPONENTS	CAS-ERM PROCESSES
Control Environment	Internal Environment	Establishing Context
	Objective Setting	
	Event Identification	Identifying Risks
Risk Assessment	Risk Assessment	Analyzing/ Quantifying Risks
	Risk Response	Integrating Risks
		Assessing/ Prioritizing Risks
Control Activities	Control Activities	Treating/ Exploiting Risks
Information and Communication	Information and Communication	
Monitoring	Monitoring	Monitoring and Reviewing

This suggests a conceptual framework as follows:

- Establishing environment/context
  - Establish management's philosophy regarding risk, recognizing that unexpected as well as expected events may occur
  - Establish the entity's risk tolerance and risk culture
  - Consider how all aspects of the entity's activities may impact the risk culture

- Setting objectives
  - Consider risk strategy in setting management objectives
  - Determine at a high level how much risk management and the board of directors are willing to accept
  - Align risk tolerance with risk appetite

- Identifying events/risks
  - Identify both internal and external occurrences that can affect strategy and achievement of objectives
  - Differentiate risks (possible negative effects) and opportunities (possible positive effects)
  - Note that a particular event may have both risk and opportunity components

- Assessing/analyzing/quantifying risks
  - Utilize both quantitative and qualitative approaches
  - Understand the extent to which events may impact objectives
  - Assess risks for both likelihood and impact

- Responding/integrating/prioritizing risks
  - Once a risk has been identified and analyzed, there are several alternatives for treating the risk:
    - Accept the risk.
      - Management "self-insures" by doing nothing
      - Accepts implications
    - Avoid the risk
      - Management eliminates the activity
    - Transfer, share, outsource the risk
      - Financial risks – Use of derivatives, hedging or insurance
      - Operational risks – Use of third parties to perform
        - Payroll processing
        - Manufacturing
        - Other back office
    - Mitigate the risk – Fix the problems
  - Evaluate the options in relation to
    - The entity's risk appetite
    - Costs vs. benefits of various responses
    - Effects of alternatives on impact and likelihood of risks
  - Select and execute the most appropriate response

- Controlling/treating risks and exploiting opportunities
  - Implement policies and procedures to ensure that management's risk tolerance and other management directives are carried out
  - Occur throughout the organization, at all levels, and in all functions
  - Include both information technology controls and application controls

- Recording, reporting, and communicating information
  - Identify, capture, and communicate relevant information in a form and on a timetable to assist stakeholders in carrying out their duties and responsibilities and evaluating opportunities
  - Communicate down, across, and up the organization

- Monitoring and reviewing
  - Conduct continuous ongoing management reviews and separate examinations to ensure the proper functioning of other ERM components
  - Adjust scope of monitoring and reviewing activities to reflect ongoing risk assessment

The elements of the process may be viewed in matrix form, as follows:

	Strategic	Operations	Financial Reporting	Compliance
Establishing environment/context	√	√	√	√
Setting objectives	√	√	√	√
Identifying events/risks	√	√	√	√
Assessing/analyzing/quantifying risks	√	√	√	√
Responding/integrating/prioritizing risks	√	√	√	√
Controlling/treating risks and exploiting opportunities	√	√	√	√
Recording, reporting, and communicating information	√	√	√	√
Monitoring and reviewing	√	√	√	√

IMPLEMENTATION GUIDANCE

The conceptual approach to implementing ERM includes the following stages:

- Planning
  - Understand the entity's environment, business model, and risk management process
  - Understand and document the entity's tone at the top and risk appetite
    - Determine risk philosophy
    - Survey risk culture
    - Consider entity's organizational integrity and ethical values

- Establish the ERM organization within the enterprise
  - Decide roles and responsibilities
  - Designate Chief Risk Officer with sufficient power to facilitate accomplishment of objectives

- Risk Assessment
  - Conduct enterprise risk assessment
    - Interviews
    - Facilitated sessions
    - Documentation
  - Train appropriate personnel for ongoing risk management activities
  - Assess risks
    - Identify
    - Measure
    - Prioritize
  - Manage risks
    - Control
    - Share or transfer
    - Diversify
    - Avoid

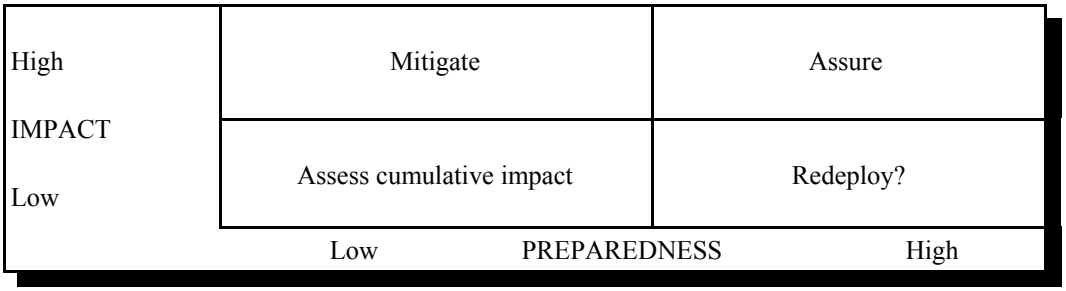
- Risk Response/Mitigation
  - Implement corrective plans/activities
  - Monitor risks
    - Process level
    - Activity level
    - Entity level
  - Monitor ongoing program development and implementation

The appropriate approach to various risk reaction and control activities depends upon the impact of the related risks and the entity's evaluation of the extent to which each of those activities prepares the entity to deal with the risk.

- If a risk has high potential impact, and the enterprise is not well prepared to handle it, immediate mitigation/remediation is required.
- If a risk has high potential impact, but the enterprise is well prepared to handle it, steps must be taken to assure that preparedness is maintained.

- If a risk has low potential impact, and the enterprise is not well prepared to handle it, immediate mitigation/remediation may not be necessary, unless a number of such risks may have a significant cumulative effect.
- If a risk has low potential impact, and the enterprise is well prepared to handle it, there is a reasonable question whether certain of the enterprise's assets and capabilities might better be redeployed to deal with more pressing risks.

This can be shown graphically as follows:



A third dimension, the likelihood that the risk will materialize, should also be considered. This may be considered in conjunction with the potential impact, so that the approach is weighted more heavily toward likely impact rather than maximum potential impact.

Common faults in implementing ERM have been found to be:

- Lack of visible, active support from Board and/or C-level management
- Implementing without a framework or plan
- Organization not ready – too much too soon
- Lack of integration with business goals and objectives
- Implementing as a project or part-time endeavor
- Failure to address the need for change management
- Failure to drive ERM to its full potential

By contrast, ERM success factors have included:

- Strong, visible support from C-level management
- Alignment of ERM to the key strategic and financial objectives and business processes
- Dedicated team of cross-functional staff to integrate ERM into significant business practices / processes
- Recognition that ERM is a continuous process and takes time to evolve

- Adequate training and supporting tools
- Leveraging well-accepted processes within the organization and introducing ERM as a value-add rather than a new stand-alone program

The changes that are required include (Wheeler & Yoo 2009):

- Clear and consistent support from Executive Management and the Board
- Long-term commitment to ERM, linked to strategic planning
- Building ERM into business processes efficiently and without undue administrative burden
- Well defined roles and responsibilities for risk, leading to improved accountability
- Risk considerations built into incentives and performance management

THOUGHTS FOR CONSIDERATION

Enterprises should consider the following thoughts with respect to their ERM effort:

- What is the number one risk facing your company today? (Wheeler & Yoo, 2009)
  - Reputational
  - Operational (technology, human capital, physical security)
  - Regulatory/legal
  - Market
  - Credit
  - Disaster (natural, terrorism)

- What is your enterprise's philosophy towards risk? (Wheeler & Yoo 2009)
  - Risk assessment
    - Annual point-in-time snapshot
    - Internal audit driven
    - Focus on current issues
  - ERM
    - Real-time, ongoing assessment
    - Continuous risk monitoring
    - Ownership of risk by process owners, embedded in the business

- How has risk management evolved in your organization? (Wheeler & Yoo 2009)
  - Developing
  - Implementing

- Improving
- Integrating

- What is the number one change barrier to overcome in your organization? (Wheeler & Yoo 2009)
  - People
    - Lack of time/skills/resources
    - Difficulty obtaining buy-in from employees
    - Lack of management support
  - Processes
    - Regulatory complexity
    - Difficulty defining risk appetite
    - Unclear lines of responsibility
  - Information
    - Lack of available data
    - Threats from unknown/unforeseeable risks
    - Difficulty in identifying emerging risks

- How can you and your group foster an ERM culture within your organization?

S&P has proposed the following questions for management meetings (S&P, 2009):

- What are the company's top risks, how big are they, and how often are they likely to occur? How often is the list of top risks updated?
- What is management doing about top risks?
- What size quarterly operating or cash loss has management and the board agreed is tolerable?
- Describe the staff responsible for risk management programs and their place in the organization chart. How do you measure success of risk management activities?
- How would a loss from a key risk impact incentive compensation of top management on planning/budgeting?
- Tell us about discussions about risk management that have taken place at the board level or among top management when making strategic decisions.
- Give an example of how your company has responded to a recent "surprise" in your industry and describe whether the surprise affected your company differently from others.

The following questions must be answered if ERM is to be made "real" for an enterprise (Baker 2008):

- What do we want to accomplish?
- What could stop us from accomplishing it?
- What should we do to make sure that those things either (1) don't happen, or (2) can be managed if they do happen?

CONCLUSION

As the complexity of modern life, and the speed with which things happens, increases continuously, the need for an effective ERM is steadily and continuously increasing. In implementing ERM, the most important consideration may be to remember what ERM is and can do, and perhaps more importantly, what it is not and cannot do.

Consistent with this discussion, ERM is about:

- Identifying and assessing key risks
- Designing and implementing processes by which those risks can be managed
- Maintaining residual risks at a level acceptable to the organization
- Linking risks back to the organizational objectives

Just as importantly, ERM is not:

- A silver bullet against bad judgment
- A once a year event
- A stand-alone, one-off initiative
- A guarantee that goals and objectives will be achieved

REFERENCES

Association of Chartered Certified Accountants (ACCA) (2008). Climbing out of the credit crunch. Retrieved May 11, 2010, from http://www.accaglobal.com/pdfs/credit_crunch.pdf

Baker, N. (2008). Real-world ERM. *Internal Auditor, 65*(6), 32-37.

Caplain, Bruce (2008). "Risk Management: Why it Failed, How to Fix It," *Internal Auditor*. Retrieved May 11, 2010 from http://www.theiia.org/intAuditor/free-feature/2008/risk-management-why-it-failed-how-to-fix-it-ii/.

Committee of Sponsoring Organizations (COSO) of the Treadway Commission (1992). *Internal control – integrated framework*. New York, NY: American Institute of Certified Public Accountants.

Committee of Sponsoring Organizations (COSO) of the Treadway Commission (2004). Enterprise risk management – integrated framework: Executive summary. Retrieved May 11, 2010, from www.coso.org/documents/COSO_ERM_ExecutiveSummary.pdf

Dreyer, S.J. & D. Ingram (2007). *Request for comment: Enterprise risk management analysis for credit ratings of nonfinancial companies*, New York, NY: Standard & Poor's (November 15).

Enterprise Risk Management Committee (2003). *Overview of enterprise risk management*. Arlington, VA: Casualty Actuarial Society (May).

Financial Stability Forum (FSF) (2009). FSF principles for sound compensation practices. (2 April). Retrieved May 11, 2010, from http://www.financialstabilityboard.org/publications/r_0904b.pdf

KPMG (2008). Value creation in a changing environment. Retrieved May 11, 2010 from http://www.kpmg.com/Global/en/IssuesAndInsights/ArticlesPublications/Pages/Value-creation-in-a-changing-risk-environment.aspx

Santori, L., K. Bevan & C. Myers (2007). *Summary of standard & poor's enterprise risk management evaluation process for insurers*. New York, NY: Standard & Poor's (November 15).

Schanfield, A. & D. Helming (2008). 12 ERM implementation challenges. *Internal Auditor, 65*(6), 41-44.

Shaw, H. (2006). The trouble with COSO: Critics say the Treadway Commission's controls framework is outdated, onerous, and overly complicated. But is there an alternative? *CFO Magazine* (March 15), retrieved on May 11, 2010, from http://www.cfo.com/printable/article.cfm/5598405

Simmons, M.R. (1997). COSO based auditing. *Internal Auditor, 54*(6) 68-72.

Standard & Poor's (2008). Enterprise risk management – rating agency view. Seminar presentation.

Standard & Poor's (2009). Discussion questions for management meetings. Seminar presentation.

Wheeler, J.A., & K.K. Yoo (2009). Enterprise risk management: what's new and what's next. Institute of Internal Auditors webinar (19 February).

A CONCEPTUAL FRAMEWORK FOR E-BANKING SERVICE QUALITY IN VIETNAM

Long Pham, New Mexico State University

ABSTRACT

Service quality is one of the key factors in determining the success or failure of e-banking. To gain and sustain competitive advantages in the rival-driven e-banking market, it is thus crucial for e-banks to understand in-depth what customers perceive to be the key dimensions of service quality and what impacts the identified dimensions have on the customers' perceived overall service quality, satisfaction, and loyalty. This paper attempts, based an extensive review of relevant literature, to provide a number of hypotheses that integrate three important constructs in the context of e-banking in Vietnam - emerging as a new potential market, such as e-service quality, e-satisfaction, and e-loyalty.

INTRODUCTION

It has been observed that the incredible growth of Internet use by individuals as well as business organizations has altered the competitive arena, which is quite unique and considerably different from that of the traditional, physical marketplace. Accordingly, the distinctive character of a virtual market has prompted companies to alter their strategies of conducting business with consumers. The banking industry is no exception. Numerous banks have already been employing the Internet as an alternative service delivery channel (Such banks are referred to as e-banks hereinafter.) to traditional ones, such as face-to-face and telephone banking, in providing their customers with a variety of financial services. It has been pointed out that the introduction of e-banking services could offer both bankers and customers diverse benefits (Broderick & Vachirapornpuk, 2002). For instance, the direct interaction between the customer, and the e-bank's Web site or employees over the Internet enables the e-bank to lower its operating and fixed costs by reducing the number of employees, branch offices, and other physical facilities while maintaining a high quality level of customer service. These cost benefits could make favorable conditions for the e-bank to provide customer services with lower fees and higher interest rates on interest bearing accounts than traditional brick-and-mortar banks (e.g., Gerlach, 2000; Jun & Cai, 2001).

Thus, in order to take advantage of this new information technology, most of the traditional banks have already invested a huge amount of money in the e-banking infrastructure and served their customers through multiple service delivery channels. This financial market change creates even

more stiff competition than ever before among e-banks. Moreover, e-banks have been facing increased challenges from nontraditional institutions, such as money management companies, securities companies, and insurance companies, erosion of product and geographic boundaries, and changes in consumers' financial awareness. This unprecedented competitive market situation presents e-bankers with severe marketing and operations challenges.

Unfortunately, although many e-banks have long centered their attention on improving their e-banking service quality, they still appear to be lagging behind their customers' ever increasing demands and expectations, and struggling with retaining and expanding their loyal customer base. Obviously, to compete successfully in such a highly competitive e-banking industry, the banks should provide customers with high quality service (Mefford, 1993). In doing so, e-banks should thoroughly understand what dimensions are utilized by customers in evaluating e-banking service quality. Then, the banks can effectively take appropriate steps to enhance their e-banking service quality, and customer satisfaction and loyalty.

Up to now, a great deal of literature has identified key dimensions of customer service quality, customer satisfaction, and customer loyalty in the setting of traditional banking, where human interactions between customers and bank employees are dominant (e.g., Baumann, Burton & Elliott, 2005; Beerli, Martin & Quintana, 2004; Calik & Balta, 2006; Ehigie, 2006; Veloutsou, Daskou & Daskou, 2004). However, very little research has addressed those issues in the banking environment, where non-human interaction is a primary service delivery and communication channel (e.g., Flavian, Tores & Guinaliu, 2004; Jabnoun & Al-Tamimi, 2003; Jun & Cai, 2001; Maenpaa, 2006; Siu & Mou, 2005).

Moreover, these studies have been primarily taken in the context of North America and Europe (Pikkarainen *et al*., 2006) and to a lesser extent in other regions including a mix of developed and developing countries, such as Singapore, Taiwan, Malaysia, and Thailand (Jaruwachirathanakul & Fink, 2005).

Little research on e-banking service quality has been implemented in countries that are emerging as new potential markets with very high economic growth rates. Among these countries is Vietnam where its economic growth rate is approximately over 8% per year and population of about 90 million (Gutman *et al*., 2006). Together with Vietnam's entry into the World Trade Organization dated on 7 January 2007, its banking sector is increasingly being deregulated in accordance with the requirements set up by the World Trade Organization. These moves would strengthen competition among local and foreign banks in Vietnam, bringing about myriad of opportunities for banks that provide superior service quality, especially e-banking service quality, for their customers.

Therefore, the objective of this research is, based on relevant literature reviews, to provide a conceptual framework that integrates e-banking service quality, customer satisfaction, and customer loyalty in the context of Vietnam. More specifically, the present study attempts to (1) identify the salient e-banking service quality dimensions; (2) examine the relationships between the

derived e-banking service quality dimensions and customer satisfaction; and (3) investigate the association between customer satisfaction and customer loyalty.

BACKGROUND

E-Service Quality

Although both academicians and practitioners appear to continuously claim about what really constitute service quality across various industries, they are increasingly reaching the consensus that service quality is determined by the difference between customers' expectations of service providers' performance and their evaluation of the services they received (Parasuraman, Zeithaml & Berry, 1985, 1988). Parasuraman, Zeithaml and Berry (1985) have originally identified ten dimensions of service quality that substantially affect the customers' perceptions of overall service quality. These determinants were tangibles, reliability, responsiveness, competence, courtesy, credibility, security, access, communication, and understanding the customer. Parasuraman, Zeithaml and Berry (1988) later refine the ten dimensions into five based on factor analysis. These five dimensions are tangibles, reliability, responsiveness, assurance, and empathy. On the grounds of these five dimensions, they have developed a 22 item survey instrument namely SERVQUAL for measuring service quality. The SERVQUAL instrument has been widely used to value the service quality of a variety of service organizations, including banks (e.g., Cowling & Newman, 1995; Jabnoun & Al-Tamimi, 2003), although it has received some criticism (for a comprehensive review, see Cronin & Taylor, 1994; Dabholkar, Thorpe & Rentz, 1996).

It is apparent that SERVQUAL may not be sufficient for measuring service quality across industries, not to mention online businesses. The instrument doest not take distinct aspects of e-service quality into consideration, since the five dimensions mainly focus on customer-to-employee, but not on customer-to-Web-site interactions. By the same token, some studies have been carried out in attempts to pinpoint major attributes that best fit the e-business setting. Cox and Dale (2001) argue that with the absence of non-human interactions in the e-setting, determinants such as competence, courtesy, cleanliness, comfort and friendliness, helpfulness, care, commitment, and flexibility were not particularly important, whereas other determinants such as accessibility, communication, credibility, understanding, appearance, and availability, were especially relevant to the success of e- businesses. Through 54 students' evaluations on three UK-based Internet bookshops, Barnes and Vidgen (2001) adjust the SERVQUAL scale and develop a WebQual Index including 24 items. This Index concentrated on seven customer service quality aspects – responsiveness, competence, reliability, access, communication, credibility, and understanding the individual.

Zeithaml, Parasuraman and Malhotra (2001), based on the traditional service quality scale and a series of focus group interviews, have developed e-service quality dimensions for measuring

e-service quality. These dimensions were access, ease of navigation, efficiency, flexibility, reliability, personalization, security/privacy, responsiveness, trust/assurance, site aesthetics, and price knowledge. Later, Wolfinbarger and Gilly (2002) rely on focus group interviews and an online survey, reduce the e-service quality scale into four main dimensions as customer service, privacy/security, reliability, and Web site design where reliability and Web site design are the most important. In addition, Madu and Madu (2002) have uncovered 15 e-service quality dimensions based on their literature review: performance, features, structure, aesthetics, reliability, storage capacity, serviceability, security and system integrity, trust, responsiveness, product differentiation and customization, Web store policies, reputation, assurance, and empathy. Moreover, Zeithaml, Parasuraman and Malhotra (2002) have proposed seven e-service quality dimensions – efficiency, reliability, fulfillment, privacy, responsiveness, compensation, and contact, in which the first four dimensions involved core e-service and the rest were relevant to service recovery.

More recently, based on focus group interviews, Santos (2003) has unfolded two groups of e-service quality dimensions that strongly affect customer retention: incubative and active groups. The dimensions of the active group are mainly related with e-consumer service quality. They consist of reliability, efficiency, support, communication, security, and incentive. Cai and Jun (2003) have come up with the following four major dimensions of e-service quality: Web site design/content, trustworthiness, prompt/reliable service, and communication. They find that all of the four dimensions substantially impact e-purchasers' evaluation of overall e-service quality. Yang, Jun and Peterson (2004) have proposed the following six e-retailer service quality dimensions: reliability, access, ease of use, attentiveness, security, and credibility. According to Lee and Lin (2005), key e-service quality dimensions are Web site design, reliability, responsiveness, trust, and personalization. They have noted that trust is the most important determinant that influences overall service quality and customer satisfaction, followed by reliability and responsiveness. In addition, Parasuraman, Zeithaml and Malhotra (2005) have developed E-S-Qual as a measure of e-core service quality, comprising four dimensions, such as efficiency, fulfillment, system availability, and privacy and E-RecS-Qual as a measure of e-recovery service quality, consisting of three dimensions, such as responsiveness, compensation, and contact.

E-Banking Service Quality

Many banks have utilized the Internet as a channel designed to offer customers a variety of financial services 24 hours a day. These services, of course, involve interactions between customers and banks' online information systems. More specifically, As noted by Rotchanakitumnuai and Speece (2003), e-banking makes favorable conditions for customers to access directly into their financial information and to make financial transactions with no need to go to the bank at any time.

Despite the importance of exploring the construct of e-banking service quality, there has been scant literature that seeks to capture salient e-banking service quality attributes. Jun and Cai

(2001) have, based on the analysis of 532 critical incidents in e-banking, developed 17 dimensions of e-banking service quality: product variety/diverse features, reliability, responsiveness, competence, courtesy, credibility, access, communication, understanding the customer, collaboration, continuous improvement, content, accuracy, ease of use, timeliness, aesthetics, and security. They suggest that both e-only banks and traditional banks offering e-banking services should focus on responsiveness, reliability, and access dimensions. Polatoglu and Ekin (2001) investigate the Turkish consumers' acceptance of e-banking service and highlight three attributes that are very likely to influence the quality of e-banking service: reliability, access, and savings.

In addition, Broderick and Vachirapornpuk (2002), employing a participant observation technique and utilizing the data of 160 incidents from 55 topic episodes posted in the bulletin board by the e-banking community, have constructed a model of perceived service quality in Internet banking. They identify the following five key elements that are regarded as central influences on perceived service quality: customer expectations of the service, the image and reputation of the service organization, aspects of the service setting, the actual service encounter, and customer participation. They further note that among these elements, service setting and customer participation have the most immediate impacts on service evaluation. Flavian, Tores and Guinaliu (2004) have uncovered four dimensions, such as access to services, services offered, security, and reputation, which are perceived to have high bearings on corporate image of e-bank and e-banking service quality. Jayawardhena (2004) has derived five quality dimensions, such as Web site interface, trust, attention, and credibility, using the modified SERVQUAL scales. Similarly, Bauer and Hammerschmidt (2005) propose a total six dimensions of e-banking portal service quality: security, trust, additional services, added values, transaction support, and responsiveness.

In addition, e-SERVQUAL was adapted by Siu and Mou (2005) in their measuring service quality in e-banking of Hong Kong. Having used factor analysis, they have unfolded four dimensions, such as credibility, efficiency, security, and problem handling. Among these four dimensions, only efficiency was found to have remained the same as the original construct and the rest were newly generated. More recently, Maenpaa (2006) has, based on open-ended exploratory interviews, an extensive literature review, and quantitative analyses, developed seven dimensions of e-banking service quality: convenience, security, status, auxiliary features, personal finances, investment, and exploration. The researcher further suggests that banks offering e-banking services need to focus more on the growing consumer cluster of youngsters, who are viewed as the prospects of tomorrow. Recently, Pikkarainen *et al*. (2006) have taken e-banking services into consideration based on an end-user computing satisfaction perspective. They strongly argue that three dimensions – content, ease of use, and accuracy - are valid in measuring end-user computing satisfaction of e-banking. Furthermore, their results elicit a solid relationship between these dimensions and overall satisfaction of e-banking.

E-Banking Services in Vietnam

According to VinaCapital (2008), there are currently four state owned commercial banks, 37 joint stock commercial banks, five joint venture banks, 28 foreign owned banks, 982 cooperatives, two policy lending banks, 55 non-bank financial institutions operating in Vietnam. There is no doubt that the number of banks has been expanding since Vietnam's entry into the World Trade Organization dated 7 January 2007. Since 1992, Vietnam has transformed its banking system into a diversified system in which commercial banks of all kinds provided services to a broader customer base. However, the state owned commercial banks account for approximately 70% of all lending activity. In 2005, foreign banks and joint ventures accounted for around 14% of lending activity. Giant foreign banks such as HSBC, Deutsche Bank and ANZ have all established their image and branches, and some have purchased shares in domestic commercial banks (VinaCapital, 2008). Most of the banks have been implementing e-banking services besides the traditional ones, for example:

- *The Bank for Foreign Trade of Vietnam (Vietcombank) started introducing its e-banking services in 2001. Its e-banking services allow customers to transfer money electronically; to get access to information such as account balance, exchange rates, and consultative information. In addition, Vietcombank's Connect 24Card allows customers to withdraw money from private accounts and international credit cards, check their account balance, make statement enquiry and transfer funds. Besides maintaining good business relationship with its long lasting customers such as state run corporations, large enterprises and import-export corporations, Vietcombank has also focused on small, medium companies and individual customers.*

- *The Industrial and Commercial Bank of Vietnam started introducing its e-banking services in 2000. This kind of service has allowed customers to get access to information such as their account balance, their recorded transactions, interest rates, exchange rates, and so on via its web-site. The bank is now co-operating with some multi-national companies, such as Fujitsu, Intel and HP to develop more complete services relating to e-banking.*

- *The Bank for Investment and Development of Vietnam (BIDV) started introducing its e-banking service in 1998. Customers can check their account balance, transfer money and pay bills. BIDV's traditional customers are enterprises operating in the fields of information technology, telecommunication, building and construction. Because BIDV is primarily operating in large cities and towns, BIDV's e-banking focuses mainly on high income and enterprise customers.*

- The Bank for Agriculture and Rural Development (Agribank) started launching its e-banking services in 2003. With a network of 1,650 branches and a number of transaction offices nationwide, Agribank has co-operated with Western Union in offering remittance services to Vietnamese overseas and migration labors in 2,800 spots throughout Vietnam.

- *Most of the other local banks and all the foreign banks operating in Vietnam have been offering e-banking services. For example, ANZ's e-banking offers customers secure and immediate e-banking services which include account balance inquiries, transaction history, funds transfer between accounts, account statement ordering, check book ordering and exchange rates.*

HYPOTHESES

Based on an extensive review of the literature on e-service quality in general and e-banking service quality in particular, the author has developed a number of hypotheses that aim at delineating the associations between e-banking service quality dimensions, overall e-banking service quality, e-banking customer satisfaction, and e-banking customer loyalty in the context of Vietnamese banking system.

E-Banking Service Quality Dimensions and Overall E-banking Service Quality

There is no doubt that to survive in the ever-increasingly competitive e-banking industry, banks need to offer customers excellent quality services. As mentioned in the review on e-banking service quality earlier, few studies have attempted to identify key dimensions of e-banking service quality and examined their relative importance to overall service quality as perceived by e-banking customers. Jun and Cai (2001) suggest that responsiveness, reliability, and accesses are the most important dimensions of e-banking service quality. According to Polatoglu and Ekin (2001), reliability, access, and savings were very likely to influence strongly the quality of e-banking service. In the view of Broderick and Vachirapornpuk (2002), service setting and customer participation were the most immediate impact on service evaluation. In addition, Flavian, Tores and Guinalie (2004) argued that access to services, services offered, security, and reputation were perceived to have high bearings on corporate image of e-bank and e-banking service quality. In the same vein, Pikkarainen *et al*. (2006) contended that the dimensions of content, ease of use, and accuracy were the most important in measuring end-user computing satisfaction of e-banking. Considering the fact that various e-banking service dimensions were uncovered by different e-banking service researchers, it would be worth validating their findings with respect to the issues of what dimensions constitute e-banking service quality and whether or not each salient e-banking

service quality dimension significantly affects customer perceived overall e-banking service quality. Therefore,

H1. Each of the dimensions of e-banking service quality will significantly influence the overall customer perceived e-banking service quality.

Overall E-Banking Service Quality and Customer Satisfaction

Banks should delight their customers by exceeding their expectations to escalate customer satisfaction (Oliver, 1980). It should be noted that the expectancy/disconfirmation paradigm in the process theory established the foundation for a significant number of satisfaction research (Mohr, 1982). This paradigm consists of four constructs as expectations, performance, disconfirmation, and satisfaction. Based on the expectancy/disconfirmation paradigm, Tse and Wilton (1988) have defined satisfaction as "the consumer's response to the evaluation of the perceived discrepancy between prior expectations and the actual performance of the product as perceived after its consumption". Seemingly, this definition is very close to that of the service quality construct. However, there are a web of distinctions between customer satisfaction and service quality. Satisfaction is a post decision customer experience, whereas quality is not (Bolton & Drew, 1991; Boulding *et al*., 1993; Cronin & Taylor, 1994; Oliver, 1980, 1993; Parasuraman, Zeithaml & Berry, 1988). Moreover, in the satisfaction literature expectations reflect anticipated performance (Churchill & Suprenent, 1982) made by the customer as to the levels of performance during a transaction. In contrast, in the service quality literature, expectations are regarded as a normative standard of future wants (Boulding *et al*., 1993). These normative standards symbol prolonged wants and needs that are kept unaffected by the adequate domain of marketing and competitive forces. Normative expectations are, hence, more stable and can be considered as representing the service the market oriented provider must constantly strive to provide (Zeithaml, Berry & Parasuraman, 1993).

There has, up to date, been a disagreement about what constitutes satisfaction. In attempts to specify the customer satisfaction construct, Giese and Cote (2000) have implemented a research that addressed a review of the satisfaction literature together with group and personal interviews. They view the customer as the final user of a product. Their study findings reveal three attributes that incorporated the construct of customer satisfaction: (1) customer satisfaction is a summary affective response that varies in intensity; (2) the response is related to a particular focus, a product choice, a purchase, or consumption; and (3) the response happens at a given time varying by circumstance, but is in general confined to time.

There has been a popular support for the proposition that customer satisfaction is an important variable in bank marketing management (Howcroft, 1991; Moutinho & Brownlie, 1989; Moutinho, 1992). The role of service quality in financial service delivery has also been spotlighted

(Avkiran, 1994; Smith & Lewis, 1989). There may be many antecedents of customer satisfaction (Jamal & Naser, 2002). However, customer satisfaction often relies much on the quality of product or service offering (Naser, Jamal & Al-Khatib, 1999). Thus, it is logical to conjecture that service quality is an antecedent to satisfaction and is non-experiential in nature (Lee, Lee & Yoo, 2000; Oliver, 1993).

In the same spirit, Caruana (2002) has examined the effects of service quality and the mediating role of customer satisfaction in the retail banking, and supports for the contention that customer satisfaction performs a mediating role in the link between service quality and service loyalty. In this study, service quality has been found to be an important input to customer satisfaction. Furthermore, Jamal and Naser (2002) argue, in the study of impact service quality dimensions and customer expertise on satisfaction in the retail banking, that the core and relational dimensions of service quality are causal antecedents of customer satisfaction. Ting (2004) also has studied service quality and satisfaction judgments of customers in banking institutions throughout Malaysia and find that service quality is the antecedent of satisfaction. Recently, Pikkarainen *et al*. (2006) have examined e-banking services and suggest that there is a positively relationship between e-banking service quality and overall satisfaction. Therefore,

H2. There is a significantly positive relationship between the overall customer e-banking service quality and e-banking customer satisfaction.

E-Banking Customer Satisfaction and Customer Loyalty

The term loyalty has been defined in a number of ways by many scholars. There are two outstanding approaches to conceptualizing the construct: behavioral and attitudinal (Dekimpe *et al*., 1997). In the behavioral approach, loyalty is elicited from customers observed purchase behavior, namely repetitive buying activity. Dick and Basu (1994) point out that the behavioral approach is inadequate to explain how and why loyalty is developed and retained and that to divulge real loyalty it is important to understand the attitudinal attributes determining repetitive purchase. Under the attitudinal approach, loyalty is hence elicited from the customer's attitude and behavioral intention towards the attitude object. These two approaches are likely to be merged by utilizing traditional attitude theory in which one of the primary premises is that behavior towards the object is determined by attitude towards the object and intention to act towards the object (Fishbein & Ajzen, 1975). More specifically, a causal chain is assumed from cognition to affect, from affect to intention, and from intention to behavior (Fishbein, 1980).

Since broadening a loyal customer base is widely accepted by academicians and practitioners as an extremely important competitive weapon to survive in today's stiff marketplace, many banks have developed and implemented diverse strategies and action programs to heighten their customer loyalty (Bahia & Nantel, 2000; Jamal & Naser, 2002). It is noteworthy that a loyal customer to a

bank is one who will stay with the same service provider, who is likely to take out new products with the bank, and who is likely to recommend the bank's services to the other people (Fisher, 2001). Among a number of factors that have been considered as significant antecedents to customer loyalty, customer satisfaction is commonly recognized by many researchers for its basic role (Jamal & Naser, 2002). Satisfied customers are more likely to focus their business with one bank (Reichheld, 1993), give recommendations for the bank and tend to decrease the bank's cost of providing services because there are fewer complaints to deal with. Moreover, Beerli, Martin and Quintana (2004) empirically investigate the factors determining e-banking customer loyalty and conclude that both satisfaction and switching costs can be regarded as loyalty antecedents and that the influence exerted by satisfaction is far greater than that of switching costs. Recently, Ehigie (2006) has conducted a study to examine how customer expectations, perceived service quality and satisfaction predict loyalty among bank customers in Nigeria. The results from this study, based on multivariate analysis, reveal that perceived service quality and customer satisfaction are jointly associated with customer loyalty, but not customer expectation. Thus, to gain customer loyalty, bank management ought to satisfy their customers. Therefore,

H3. There is a significantly positive relationship between e-banking customer satisfaction and e-banking customer loyalty.

CONCLUSION

With the Internet and Web technologies, e-banking customers can have unlimited access to the information they require and enjoy a wider range of choices in selecting banking products and services with highly competitive prices. As a result, it is generally difficult for e-banks to gain and sustain competitive advantages based solely on a cost leadership strategy in the rival-driven online banking market (Jun, Yang & Kim, 2004).

Therefore, the service quality levels of the e-banks have increasingly become a key driving force in enhancing customers' satisfaction and in turn expanding their loyal customer bases. Service quality improvement initiatives should begin with defining the customers' needs and preferences, and their related quality dimensions. By understanding the dimensions that customers use to evaluate service quality, the e-banks can take appropriate actions to monitor and enhance their performance on these dimensions. Since few studies have examined systemically the relationships between e-banking service quality dimensions, overall e-banking service quality, customer satisfaction, and customer loyalty, the author of this study, to fill this research gap, have proposed a number of hypotheses in which the aforementioned constructs are integrated in the context of Vietnamese banking system.

REFERENCES

Avkiran, N.K (1994). Developing an instrument to measure customer service quality in branch banking. *International Journal of Bank Marketing, 12*(6), 10 - 18.

Bahia, K & J. Nantel (2000). A reliable and valid measurement scale for the perceived service quality of banks. *International Journal of Bank Marketing, 18*(2), 84 - 91.

Barnes, S.J & R. Vidgen (2001). An evaluation of cyber-bookshops: The WebQual method. *International Journal of Electronic Commerce, 6*(1), 11 - 30.

Bauer, H.H., M. Hammerschmidt & T. Falk (2005). Measuring the quality of e-banking portals. *International Journal of Bank Marketing, 23*(2), 153 - 175.

Baumann, C., S. Burton & G. Elliott (2005). Determinants of customer loyalty and share of wallet in retail banking. *Journal of Financial Services Marketing, 9*(3), 231 - 248.

Beerli, A., J.D. Martin & A. Quintana (2004). A model of customer loyalty in the retail banking market. *European Journal of Marketing, 38*(1/2), 253 - 275.

Bolton, R.N & J.H. Drew (1991). A multistage model of customers' assessments of service quality and value. *Journal of Consumer Research, 17*March, 375 - 384.

Boulding, W., A. Kalra., R. Staelin & V. Zeithaml (1993). A dynamic process model of service quality: From expectations to behavioral intentions. *Journal of Marketing Research, 30* February, 7 - 27.

Broderick, A.J & S. Vachirapornpuk (2002). Service quality in Internet banking: The importance of customer role. *Marketing Intelligence & Planning, 20*(6), 327 - 335.

Cai, S & M. Jun (2003). Internet users' perceptions of online service quality: A comparison of online buyers and information searchers. *Managing Service Quality, 13*(6), 504 - 519.

Calik, N & N.F. Balta (2006). Consumer satisfaction and loyalty derived from the perceived quality of individual banking services: A field study in Eskisehir from Turkey. *Journal of Financial Services Marketing, 10*(4), 135 - 149.

Caruana, A (2002). Service loyalty: The effects of service quality and the mediating role of customer satisfaction. *European Journal of Marketing, 36*(7/8), 811 - 828.

Churchill, G.A.Jr & C. Surprenant (1982). An investigation into the determinants of customer satisfaction. *Journal of Marketing Research,* (21)November, 491 - 504.

Cowling, A & K. Newman (1995). Banking on people: TQM, service quality, and human resources. *Personnel Review, 24*(7), 25 - 40.

Cox, J & B.G. Dale (2001). Service quality and e-commerce: An exploratory analysis. *Managing Service Quality, 11*(2), 121 - 131.

Cronin, J.J & S.A. Taylor (1994). SERVPERF versus SERVQUAL: Reconciling performance -based and perceptions-minus-expectations measurement of service quality. *Journal of Marketing, 58*(1), 125 - 131.

Dabholkar, P.A., D.I. Thorpe & J.O. Rentz (1996). A measure of service quality for retailing stores: Scale development and validation. *Journal of the Academy of Marketing Science, 24*(1), 2 - 16.

Dekimpe, M.G., J.B.E.M. Steenkamp., M. Mellens & P.V. Abeele (1997). Decline and variability in brand loyalty. *International Journal of Research in Marketing, 14*, 405 - 420.

Dick, A.S & K. Basu (1994). Customer loyalty: Toward an integrated conceptual framework. *Journal of the Academy of Marketing Science, 22*(2), 99 - 113.

Ehigie, B.O (2006). Correlates of customer loyalty to their bank: A case study in Nigeria. *International Journal of Bank Marketing, 24*(7), 494 - 508.

Fishbein, M (1980). An overview of the attitude construct. In Hafer, G.B., A look back, a Look ahead. *American Marketing Association*, Chicago, IL.

Fishbein, M & I. Ajzen (1975). *Belief, attitude, intention, and behavior.* Addison-Wesley, Reading, MA.

Fisher, A (2001). Winning the battle for customers. *Journal of Financial Service Marketing, 6*(2), 77 - 83.

Flavian, C., E. Tores & M. Guinaliu (2004). Corporate image measurement: A further problem for the tangibilization of Internet banking services. *The International Journal of Bank Marketing, 22*(5), 366 - 384.

Gerlach, D (2000). Put your money where your mouse is. *PC World* March, 191 - 199.

Giese, J & J. Cote (2000). Defining customer satisfaction. *Academy of Marketing Science Review,* from http://www.amsreview.org/amsrev/theory/giese00-01.html

Howcroft, B (1991). Customer service in selected branches of a UK clearing bank: A pilot study. *Proceedings of the Service Industries Management Research Unit Conference.* Cardiff Business School, University of Wales College of Cardiff, 25-26 September.

Jabnoun, N & H.A.H. Al-Tamimi (2003). Measuring perceived service quality at UAE commercial banks. *International Journal of Quality & Reliability Management, 20*(4), 458 - 472.

Jamal, A & K. Naser (2002). Customer satisfaction and retail banking: An assessment of some of the key antecedents of customer satisfaction in retail banking. *International Journal of Bank Marketing, 20*(4/5), 146 - 161.

Jaruwachirathanakul, B & D. Fink (2005). Internet banking adoption strategies for a developing country: the case of Thailand. *Internet Research, 15*(3): 295 – 311.

Jayawardhena, C (2004). Measurement of service quality in Internet banking: The development of an instrument. *Journal of Marketing Management, 20*, 185 - 207.

Jun, M & S. Cai (2001). The key determinants of Internet banking service quality: A content analysis. *International Journal of Bank Marketing,* July, 276 - 291.

Jun, M., Z. Yang & D.S. Kim (2004). Customers' perceptions of online retailing service quality and their satisfaction. *International Journal of Quality & Reliability Management, 21*(8), 817 - 840.

Lee, G.G & H.F. Lin (2005). Customer perceptions of e-service quality in online shopping. *International Journal of Retail & Distribution Management, 33*(2), 161 - 176.

Lee, H., Y. Lee & D. Yoo (2000). The determinants of perceived service quality and its relationship with satisfaction. *Journal of Services Marketing, 14*(3), 217 - 231.

Madu, C.N & A.A. Madu (2002). Dimensions of e-quality. *International Journal of Quality & Reliability Management, 19*(3), 246 - 258.

Maenpaa, K (2006). Clustering the consumers on the basis of their perceptions of the Internet banking services. *Internet Research, 16*(3), 304 - 322.

Mefford, R.N (1993). Improving service quality: Learning from manufacturing. *International Journal of production Economics, 30*, 399 - 413.

Mohr, L.B (1982). *Explaining organizational behavior*. Jossey-Bass, San Francisco, CA.

Moutinho, L (1992). Customer satisfaction measurement: Prolonged satisfaction with ATMs. *International Journal of Bank Marketing, 10*(7), 30 - 37.

Moutinho, L & D.T. Brownlie (1989). Customer satisfaction with bank services: A multidimensional space analysis. *International journal of Bank Marketing, 7*(5), 23 - 27.

Naser, K., A. Jamal & K. Al-Khatib (1999). Islamic banking: A study of customer satisfaction and preferences in Jordan. *International Journal of Bank Marketing, 17*(3), 135 - 150.

Oliver, R.L (1980). Cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research, 17*(11), 460 - 469.

Oliver, R.L (1993). A conceptual model of service quality and service satisfaction: Compatible goals, different concepts. *Advances in Services Marketing and Management, 2*, 65 - 85.

Parasuraman, A., V.A. Zeithaml & L.L. Berry (1985). A conceptual model of service quality and its implications for future research. *Journal of Marketing, 49*(4), 41 - 50.

Parasuraman, A., V.A. Zeithaml & L.L. Berry (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing, 64*(1), 12 - 40.

Parasuraman, A., V.A. Zeithaml & A. Malhotra (2005). E-S-Qual: A multiple-item scale for assessing electronic service quality. *Journal of Service Research, 7*(3), 213 - 233.

Pikkarainen, K., T. Pikkarainen., H. Karjaluoto & S. Pahnila (2006). The measurement of end-user computing satisfaction of online banking services: Empirical evidence from Finland. *International Journal of Bank Marketing, 24*(3), 158 - 172.

Polatoglu, V.N & S. Ekin (2001). An empirical investigation of the Turkish consumers' acceptance of Internet banking services. *International Journal of Bank Marketing,* April, 156 - 165.

Reichheld, F (1993). Loyalty based management. *Harvard Business Review,* March-April, 64 - 73.

Rotchanakitumnuai, S & M. Speece (2003). Barriers to Internet banking adoption: A qualitative study among corporate customers in Thailand. *International Journal of Bank Marketing, 21*(6), 312 - 323.

Santos, J (2003). E-service quality: A model of virtual service quality dimensions. *Managing Service Quality, 13*(3), 233 - 246.

Siu, N.Y.M & J.C.W. Mou (2005). Measuring service quality in Internet banking: The case of Hong Kong. *Journal of International Consumer Marketing, 17*(4), 99 - 116.

Smith, A.M & B.R. Lewis (1989). Customer care in financial service organizations. *International Journal of Bank Marketing, 7*(5), 13 - 22.

Ting, D.H (2004). Service quality and satisfaction perceptions: Curvilinear and interaction effect. *The International Journal of Bank Marketing, 22*(6), 407 - 420.

Tse, D.K & P.C. Wilton (1988). Models of consumer satisfaction formation: An extension. *Journal of Marketing Research, 17* November, 460 - 469.

Veloutsou, C., S. Daskou & A. Daskou (2004). Analysis papers: Are the determinants of bank loyalty brand specific. *Journal of Financial Services Marketing, 9*(2), 113 - 125.

VinaCapital (2008). Retrieved August 10, 2009, from http://www.vinacapital.com.

Wolfinbarger, M.F & M.C. Gilly (2002). ETailQ: Dimensionalization, measuring and predicting Etail quality. *Journal of Retailing, 79*(3), 183 - 198.

Yang, Z., M. Jun & R.T. Peterson (2004). Measuring customer perceived online service quality: Scale development and managerial Implications. *International Journal of Operations & Production Management, 24*(11), 1149 - 1174.

Zeithaml, V.A., L.L. Berry & A. Parasuraman (1993). The nature and determinants of customer expectations of service. *Journal of the Academy of Marketing Science, 21*(1), 1 - 12.

Zeithaml, V.A., A. Parasuraman & A. Malhotra (2001). A conceptual framework for understanding e-service quality: Implications for future research and managerial practice. *MSI Working Paper Series* Report No. 00-115, Cambridge, MA.

Zeithaml, V.A., A. Parasuraman & A. Malhotra (2002). Service quality delivery through web sites: A critical review of extant knowledge. *Journal of the Academy of Marketing Science, 30*(4), 362 - 375.

SPOKES-CHARACTER OF THE NATION'S FIRST STATEWIDE BOOSTER SEAT SAFETY PROGRAM: OLLIE OTTER SAFETY MASCOT

Amanda L. Brown, Tennessee Tech University
Ismet Anitsal, Tennessee Tech University
M. Meral Anitsal, Tennessee Tech University
Kevin Liska, Tennessee Tech University

ABSTRACT

Increased awareness of the importance of elementary school children's using booster seats is necessary to keep children safe while riding in motor vehicles. One effective teaching tool in heightening such awareness is a spokes-character used in commercials, in public service announcements, and in product packaging. The Ollie Otter Booster Seat and Seat Belt Program features a spokes-character, Ollie Otter, to promote booster seats and seat belts to Tennessee's elementary school students, who are vulnerable to head and other injuries if child restraint systems are discontinued too soon. This visually oriented program, supported through multiple sponsorships and partnerships, has experienced phenomenal growth since its inception. This paper examines the program's reach (including serving as a prototype for other states), a spokes-character's ability to increase awareness of and interest in using booster seats among K-4 children, and areas for future research.

INTRODUCTION

Of growing concern for Americans has been the alarming number of children killed in motor vehicle accidents. Rice and Anderson (2009) stated, "Motor vehicle collisions are the leading cause of unintentional injury and death among children aged 1 year and older in the United States". As a result, several different steps have been taken to keep children safer while riding in a vehicle. Durbin et al (2003) discovered "that the odds of serious injury were fifty-nine percent lower for crash-involved children aged 4 to 7 years using booster seats and lap-shoulder belts compared with lap-shoulder belts only" as cited in Miller et al (2006, p.1995). According to The National Highway Traffic Safety Administration, belt positioning booster seats (BPBs) are the recommended restraint for 4- to 8-year-olds (Philbrook et al 2009). Research done by Winston et al revealed, "Premature graduation of young children from child restraint systems (CRS) to seat belts puts them at greatly

increased risk of significant injury in crashes. A major benefit of CRS is a reduction in head injuries, potentially attributable to a reduction in the amount of head excursion in a crash" (2000, p. 1183). Booster seats make seat belts more comfortable and may make children more willing to sit in the seat, thus improving behavior in the vehicle (Miller et al 2006).

Many methods have been used to increase awareness of booster seats' necessity. For example, laws have been implemented in thirty-six states, and law-enforcement officials nationwide supervise checkpoints to encourage properly using child-restraint systems. Physicians have also been encouraged to speak to parents about the consequences of not restraining their children. However, 86 percent of children who should be restrained in car seats or belt-positioning booster seats are inappropriately placed in seat belts (Simpson, et al 2002).

Many parents refer to state laws in making decisions about implementing passenger-safety devices for children. While these laws provide mandatory information, they do not, however, always include all the information necessary for optimal restraint or the "best practices" for questionable situations. According to Simpson et al, although child passenger safety laws are improving, parents should be cautioned against using this source [these laws] as the sole determinant of child restraint choice (2002).

Furthermore, while information is readily available, some parents are unaware of the laws for using booster seats and seat belts. Testing the effectiveness of different techniques, a study in twenty-four elementary schools found that written information alone was not effective in providing booster seat education. Instead, "Providing instruction to parent groups and teaching children in the classroom about booster seats were shown to improve booster seat use" (Philbrook et al 2009, p.220). Heightening awareness, The Ollie Otter Booster Seat and Seat Belt Safety Program uses a spokes-character, Ollie Otter, to spread this information to children in elementary schools. For younger children, the program focuses on using booster seats; for older children, the focus changes to seat belt use and positions for younger children.

This paper explores the frequent use of spokes-characters in television advertising specifically targeted to children. Also discussed is the need for further education among elementary school children concerning promotion of a positive brand image and the increased use of booster seats. A final consideration is the use of a spokes-character by a nonprofit organization, the Ollie Otter Booster Seat and Seat Belt Safety Program, and how using this character has led to the program's growth.

USE OF SPOKES-CHARACTERS

The Ollie Otter Booster Seat and Seat Belt Program is a "brand" that is marketing a product, "child passenger safety," using spokes-character Ollie Otter to appeal to consumers, "children." Phillips and Gyoerick (1999) define *spokes-character* as "an animate being or animated object that is used to promote a product, service, or idea" (p.714). These researchers emphasize that a "spokes-

character does not have to be a legal trademark or appear on the package but must be used consistently in conjunction with a product over time" (p. 714).

Examples of spokes-characters range from animated characters on television to furry mascots in gymnasiums. "Many of today's popular characters, such as the Jolly Green Giant, Betty Crocker, and Mr. Peanut, have been used consistently by advertisers for more than seventy years" (Phillips and Gyoerick 1999, p.713). Animal personifications have been popular over the years and continue to be used because of their appeal (Callcott and Lee 1994). One notable example is Smokey the Bear with his motto, "Only you can prevent wild fires." Created in 1941 by the Ad Council, the Smokey the Bear campaign was the longest-running public-service-announcement campaign in United States' history (www. Smokeybear.com). Spokes-characters bring many positive traits to a product, specifically for children. Honesty, trust, and expertise are a few characteristics spokes-characters convey (Garretson and Niedrich 2004). The choices these characters model have enormous influence on consumers (LeBel and Cooke 2008). For example if children view Ollie Otter as an expert on child passenger safety, they are more likely to be attentive during the presentation and remember the important facts.

In addition, spokes-characters, such as Ollie, are not prone to the negative publicity that can come with human celebrities, who can ultimately have a bad effect on the brand or the idea that the spokesperson is representing. For example, Ollie Otter always wears his seat belt and is usually carrying a height chart, emphasizing the four-foot-nine requirement for wearing a seat belt. He models good decisions in hopes that children, will do likewise. In contrast, Stafford et al (2002) depict the possible downfalls to which a human celebrity spokesperson can fall prey while endorsing a product, brand, or idea: "The spokesperson may become a direct representation of the particular service being advertised, and this spokesperson's physical and intellectual characteristics are likely to have a bearing on how well the audience accepts the proffered claims or endorsements" (p.17). According to Crutchfield and Grant (2008), high-impact nonprofit organizations create social change by using new models. These organizations have the highest level of social impact because they are innovative and entrepreneurial (Kelly and Lewis 2009). The Ollie Otter program has successfully used its spokes-character as an innovative tool in elementary schools to distribute a product—the message of child passenger safety—to K-4 children Schools are in a prime situation to educate children on making better health decisions. As stated by Marks, "School-based health education can help young people develop the knowledge, skills, motivation, and support they need to choose health-enhancing behaviors and to resist behaviors that put them at risk for health and social problems and school failure" (2009, pg. 6).

As consumers, children tend to be more visual than verbal when retaining facts (McNeal and Ji 2003). At an early age, children can recognize characters to which they have been exposed several times and can show a desire for the character and for the products associated with that character. Therefore, a character's persuasive messages can make children knowledgeable consumers (Neeley and Schumann 2004). While advertising directly to children is an ethical issue, in the case of

nonprofit organizations, the advertising is sending a positive message about a certain product—whether child-passenger safety, a non-smoking campaign, or forest-fire prevention.

OLLIE OTTER BOOSTER SEAT PROGRAM

In December of 2006, the Ollie Otter Booster Seat and Seat Belt Safety Program (Ollie Otter Program) began in Tennessee. This program's primary objective is to offer booster seat and seat belt education to children K-4th grade across Tennessee through interactive programs in schools. The Ollie Otter program has several sponsors including:

- Tennessee Road Builders Association (TRBA), whose motto is "Good Roads Save Time, Money & Lives"
- Governor's Highway Safety Office, a division of the Tennessee Department of Transportation and Tennessee's advocate for highway safety (LINK"http://www.tdot.state.tn.us/ghso"www.tdot.state.tn.us/ghso)
- Tennessee Tech University (TTU), which The Princeton Review consistently identifies as one of the best in the Southeast (www.tntech.edu).

The program has also partnered with several Tennessee organizations including the following:

- Office of Coordinated School Health, whose primary goal is improving student-health outcomes (www.state.tn.us/education/schoolhealth/)
- Tennessee Highway Patrol, which provides education about and enforcement of all federal and state laws relating to traffic (www.state.tn.us/safety/thp.htm)
- SAFEKids, which is devoted to the prevention of unintentional childhood injury
- Tennessee Technology Centers (TTCs), which are located across Tennessee and are the state's premier providers of workforce development (www.tbr.edu/schools).

Several methods are used to implement this statewide program. The Tennessee Board of Regents (TBR) houses an online class, produced at Tennessee Tech University, to train volunteers on how to give a school presentation, how to write a press release, and how to get others involved in their communities. The Regents Online Continuing Education (ROCE) program, a division of TBR, has benefited from this partnership through increased exposure to the target market—teachers. The 725 documented press releases contain information about the online course that ROCE presents on the Ollie Otter program's behalf. In addition, an informative website (www.seatbeltvolunteer.org) is database-driven and coordinates volunteers and scheduling of school

events. This website was built and is maintained at Tennessee Tech University's BusinessMedia Center. Supplementing the program is informational material, including height charts, newsletters, volunteer posters, and a statewide public service announcement.

Over the past three years, the number of children hearing the Ollie Otter message about seat belt and booster seat safety has grown tremendously and is expected to continue growing as reflected in Table 1.

Table 1: Ollie Otter Program's Growth in Tennessee

Grant Year	Schools	Classrooms	Children
Year One	154	2,928	57,184
Year Two	312	5,037	91,500
Year Three (As of February 23, 2010)	149	2,266	41,519
Year Three (Projections)	330	5,940	118,800
Three-Year Total	**796**	**13,905**	**267,484**

Tennessee Tech's BusinessMedia Center's objective was to saturate Tennessee during the first program year—an unheard of and almost unrealistic task of any safety program. From the program's first to second year, the number of schools visited more than doubled. As of February, 23, 2010, Ollie Otter program has visited 615 elementary schools, 10,000 classrooms, and over 190,000 children. The program is expected reach its 200,000th child by the end of May 2010. To support this program's magnitude, over 700 volunteers have dedicated their time and effort. With all volunteer teams using all twelve Ollie suits in circulation, the program can be presented in 180 places during a single school week.

A day in the life of Ollie can consist of visiting an elementary school; participating in community events including county health fairs, National Child Passenger Safety Week, and seat belt and booster seat checkpoints; or recruiting new volunteers in conferences across the country, Though initiated in Tennessee, this unique program is a role model, having spurred interest in sixteen other states, provides guidance and access to materials for duplication. Mississippi has already adopted the program and has received funding to implement the program statewide. Thirty-four other states have a booster seat law and will probably soon adopt a comprehensive program such as Ollie Otter. Because of its widespread influence, the program has received numerous awards, including the American Road and Transportation Builders Association Award, the Horizon Award for Website Development, and the Lifesavers Award from the Governor's Highway Safety Office.

OLLIE OTTER: THE SPOKES-CHARACTER

As pictured in Figure 1, Ollie Otter is a furry otter mascot that conveys the message of seat belt, booster seat, and work zone safety.

Figure 1: Ollie Otter Spokes-character

Wearing a bright-orange jersey, Ollie Otter promotes safety from his head to furry toe. Orange construction barrels are used in the presentation to teach school children the importance of work-zone safety. Ollie always wears his seatbelt, which comes over his shoulder and buckles at his waist, and sports a TRBA logo on the right side of his chest and a Tennessee Highway Patrol badge on his right shoulder. Six- to seven-feet tall, depending on the volunteer wearing the costume, Ollie dominates a room (Brewer 2010).

Since the program's inception, Ollie has been used to interact and share with elementary children the importance of sitting in a booster seat and bucking up every time they get in a vehicle. He has become a celebrity in the schools he visits and always leaves a lasting impression on the children he meets.

Ollie was a dream turned reality for creator Carol Coleman, then President of the TRBA Women's' Auxiliary. After losing several family members in vehicular collisions across Tennessee, Coleman, a lifelong resident of Livingston, Tennessee, wanted to make a difference. Carol reached out to Tennessee Tech University's BusinessMedia Center in the College of Business to start a campaign teaching elementary school children the importance of safety when riding in a vehicle. Knowing that this program helps to save lives, Coleman believes that elementary school children are at "the most important age to teach them road safety, work zone safety, seatbelt and booster seat safety" (2010). Jumping at the opportunity to participate, the BusinessMedia Center is responsible for managing and maintaining the program. The Governor's Highway Safety Office provided a grant to fund the program. The name *Ollie* was chosen when Coleman asked the members of the

Tennessee Road Builders Association to ask their children and grandchildren what name they would find suitable for an otter. Although multiple responses were given, three children across Tennessee responded "Ollie." Coleman agreed that Ollie the Otter "seemed like the perfect name" (2010).

At the end of a presentation, teachers are encouraged to have the children write Ollie a letter or draw him a picture so that his message stays with them longer. Hundreds of children have shared their thoughts and words with Ollie by sending him drawings and letters. According to Julie Brewer, Ollie Otter program manager, "Ollie just has an educational message with him that is life saving and can really impact the way kids think about their safety and how they get into cars and buckle up" (2010). While other nonprofit organizations use different methods to have impact on a community, Ollie Otter is perfect for this message. By teaching young children the importance of safety, the Ollie Otter program hopes they will carry that lesson with them throughout life. Ollie brings life to the following characteristics:

Ollie is a relatable otter. Only three years old, Ollie learns new lessons everyday just like the children in the schools he visits. He listens attentively to the presentation speaker and participates in games during the presentation. Ollie is as much a kid as the ones he visits. A third grader wrote, "You are so funny. In fact, your [sic] really funny. Thanks for the bookmarks. There [sic] so cool."

Ollie is a cool otter. He struts when he walks into a room and loves to dance. Ollie conveys the message that little kids sit in a car seat and big kids get to sit in booster seats. He makes sitting in a booster seat cool, whereas most children think they are too big to sit in one. "I think Ollie brings to the children this fairy-tale childlike feeling that 'Hey guys! It is cool! It's a great thing!" says Coleman (2010). One third-grader wrote, "Thank you for saying booster seats are cool for small people. And thank you for bringing out the ruler to measure us."

Ollie is a friendly otter. At the end of every presentation Ollie stands at the gym doors and gives every child the opportunity to give him a high-five or a hug. Ollie loves to have his picture taken with the kids and loves to receive mail from them. One first grader responded, "I learned to ride in a booster and fasten my seat belt. I loved the high fives!"

Ollie is an informative otter. His ultimate goal is to bring a very important message to every school or community event where he appears. A second grader wrote, "I learned that you have to ride in a booster seat until you are 4 foot tall and 9 inches"; a third grader wrote, "I bet many people don't know that you have to be 4'9" to ride in a regular seat."

Ollie is becoming a media star. WCTE, a PBS affiliate station, aired a full episode of "Focus On" concerning the Ollie Otter program and its implementation in Tennessee. The Tennessee Road Builders Association partnered with Tennessee Highway Patrol and the Ollie Otter program to create a public service announcement that airs on statewide television stations. Ollie's fame has also inspired creativity. John Farrell, a singer and songwriter from Hillsdale, New York, wrote a song about Ollie Otter and his safety message. The following is the song's chorus:

> We're going to buckle our seat belts every time.
> We're going to take good care of ourselves.
> We're going to ask our families and our friends.
> Please, Please buckle up your belts.

We're gonna use our booster seats when we need to.
It's the law until we're four feet nine.
We're gonna ask the drivers behind the wheels
To please please please, please please please,
Drive safely all the time!

Ollie not only brings joy to children, but also influences every person who has the honor of meeting him. According to Melissa Roberson, Ollie Otter East Tennessee program coordinator, "Working with the Ollie Otter program has allowed me to gain a better understanding for the need for roadway, booster seat, and seat belt safety to be taught in schools. Volunteering for this program is a great way to give back to your community and also help save children's lives" (2010). A positive feeling is created when an individual can help make a difference in the children of an entire elementary school. Therefore, the Ollie Otter program inspires everyone involved.

CONCLUSION

Using a spokes-character to reach children, the Ollie Otter Booster Seat and Seat Belt Safety Program has a vast impact. As McNeal and Ji (2003) indicated, when children drew pictures of Ollie, they included details they found valuable, such as being larger than life and always buckled up, they also included slogans like "Buckling is cool!" These researchers also suggest that spokes-characters such as Ollie are more memorable than real people. Letters and drawings to Ollie weeks after the school visit indicate a longer retention of Ollie's identity and message in the visual and verbal memory of K-4 children.

While using a spokes-character has been effective in this case, it is unknown if such use would be effective with other nonprofit organizations geared toward children. If in fact spokes-characters are successful in other nonprofit organizations, their implementation should increase. Hopefully, this paper will heighten awareness that spokes-characters are not merely animated characters selling breakfast cereal, but can actually be a successful tool to help keep children safe.

Children show their appreciation of Ollie's safety expertise. Letters indicate that they trust Ollie knows best about child safety in vehicles. Although Ollie's conversational skills are very limited, his gestures (high fives and hugs) and body language (attentive listening of safety lecture) make him likable, appropriate and believable for children. The Ollie Otter campaign provides evidence that spokes-characters can be effective in promoting child safety among children by making buckling and booster seats "cool." This campaign's results are in line with other research. For example, Luo et al.'s (2006) findings indicate that cartoon-like characters contribute more positively to persuasiveness than human-like characters. Furthermore, Stafford, Stafford, and Day (2002) suggest that spokes-characters can generate much awareness and influence affective components of attitudes positively. Animated spokes-characters are frequently not only used in advertising for-profit products and services but also debated in advertising literature. Enhanced recognition and liking are well documented; however, their impacts on intention and product choice among young children are inconclusive (Neeley and Schumann 2004). Furthermore, research on

spokes-characters' effectiveness in the context of nonprofit organizations has not been done to the knowledge of authors yet.

This campaign's primary aim is to encourage the proper use of seat belts and booster seats within Tennessee. The next campaign wave should focus on increasing retention of safety messages among K-4 children. This campaign also serves as a catalyst for future research. For example, while spokes-characters have clearly been effective with consumers of all ages, specifically children, research on the effectiveness of spokes-characters for nonprofit organizations geared toward children has yet to be done. Several other research avenues are also available, including further investigating the retention of Ollie's message as children grow from year to year, a behavior change in parents toward the use of booster seats, and whether children have any influence on their parents' decision to buckle up and to practice other safe driving habits.

REFERENCES

Brewer, Julie (2010), Personal Interview.

Callcott, Margaret F. and Wei-Na Lee (1994). A Content Analysis of Animation and Animated Spokes-Characters in Television Commercials. *Journal of Advertising*, 23(4), 1-12.

Chandler, Tomasita M. and Barbara M. Heinzerling (1998). Learning the Consumer Role: Children as Consumers. *Reference Services Review*, 26(1), 61-95.

Coleman, Carol (2010), Personal Interview.

Garretson, Judith A. and Ronald W. Niedrich (2004). Spokes-Characters: Creating Character Trust and Positive Brand Attributes. *Journal of Advertising*, 33(2 Summer), 25-36.

Governor's Highway Safety Office (www.tdot.state.tn.us/ghso). Retrieved February 12, 2010.

Hsu, Chung-kue and Daniella McDonald (2002). An Examination on Multiple Celebrity Endorsers in Advertising. *Journal of Product & Brand Management*, 11(1), 19-29.

Kelly, Deborah and Alfred Lewis (2009). Human Service Sector Nonprofit Organization's Social Impact. *Business Strategy Series*, 10(6), 374-382.

LeBel, Jordan L. and Nathalie Cooke (2008). Branded Food Spokes-Characters: Consumers' Contributions to the Narrative of Commerce. *Journal of Product & Brand Management*, 17(3), 143-153.

Luo, J.T., Peter McGoldrick, Susan Beatty, and Kathleen A. Keeling (2006). On-Screen Characters: Their Design and Influence on Consumer Trust. *Journal of Services Marketing,* 20(2), 112-124.

Marks, Ray (2009). Schools and Health Education: What Works, What is Needed, and Why?. *Journal of Health Education*, 109(1), 4-8.

McNeal, James U. and Mindy F. Ji (2003). Children's Visual Memory of Packaging. *Journal of Consumer Marketing*, 20(5), 400-427.

Miller, Ted R., Eduard Zaloshnja, and Delia Hendrie (2006). Cost-Outcome Analysis of Booster Seats for Auto Occupants Aged 4 to 7 years. *Pediatrics*, 118(5), 1994-1998.

Neely, Sabrina M. and David Schumann (2004). Using Animated Spokes-Characters in Advertising to Young Children: Does Increasing Attention to Advertising Necessarily Lead to Product Preference?. *Journal of Advertising*, 33(3Fall), 7-23.

Office of Coordinated School Health www.state.tn.us/education/schoolhealth/ Retrieved February 21, 2010.

Philbrook, Julie K., Andrew W. Kiragu, Joni S. Geppert, Patricia R. Graham, Laura M. Richardson, and Robert L. Kriel (2009). Pediatric Injury Prevention: Methods of Booster Seat Education. *Pediatric Nursing*, 35(4), 215-220.

Phillips, Barbara J. and Barbara Gyoerick (1999). The Cow, The Cook, and the Quaker: Fifty Years of Spokes-Character Advertising. *Journalism and Mass Communication Quarterly*, 76(4 Winter), 713-728.

Rice, Thomas M. and Craig L. Anderson (2009). The Effectiveness of Child Restraint Systems for Children Aged 3 Years or Younger During Motor Vehicle Collisions: 1996 to 2005. *American Journal of Public Health*, 99(2), 252-257.

Roberson, Melissa (2010), Personal Interview.

Simpson, Edith M., Elisa K. Moll, Nancy Kassam-Adams, Gwenyth J. Miller, and Flaura K. Winston (2002). Barriers to Booster Seat Use and Strategies to Increase Their Use. *Pediatrics*, 110(4), 729-736.

Smokey the Bear www.smokeybear.com, Retrieved February 23, 2010.

Stafford, Marla Royne, Thomas F. Stafford, and Ellen Day (2002). A Contingency Approach: The Effects of Spokesperson Type and Service Type on Service Advertising Perceptions. *Journal of Advertising*, 31(2 Summer), 17-34.

Tennessee Highway Patrol www.state.tn.us/safety/thp.htm, Retrieved February 15, 2010.

Tennessee Technology Centers http://www.tbr.edu/schools/default.aspx?id=2654 Retrieved February 24, 2010.

Tennessee Technological University www.tntech.edu Retrieved February 24, 2010.

Winston, Flaura K., Dennis R. Durbin, Michael J. Kallan, and Elisa K. Moll (2000). The Danger of Premature Graduation to Seat Belts for Young Children. *Pediatrics*, 105(8), 1179-1183.