Volume 9, Numbers 1 and 2 **ISSN 1939-2230**

ACADEMY OF BANKING STUDIES JOURNAL

Kurt Jesswein
Sam Houston State University
Editor

The *Academy of Banking Studies Journal*, formerly known as the *Journal of Commercial Banking and Finance*, is owned and published by the DreamCatchers Group, LLC, and printed by Whitney Press, Inc.

Whitney Press, Inc.

Printed by Whitney Press, Inc.
PO Box 1064, Cullowhee, NC 28723
www.whitneypress.com

All authors must execute a publication permission agreement taking sole responsibility for the information in the manuscript. The DreamCatchers Group, LLC is not responsible for the content of any individual manuscripts. Any omissions or errors are the sole responsibility of the individual authors.

The *Academy of Banking Studies Journal* is owned and published by the DreamCatchers Group, LLC, PO Box 2689, 145 Travis Road, Cullowhee, NC 28723, U.S.A., (828) 293-9151, FAX (828) 293-9407. Those interested in subscribing to the *Journal*, advertising in the *Journal*, submitting manuscripts to the *Journal*, or otherwise communicating with the *Journal*, should contact the Executive Director of the Allied Academies at info@alliedacademies.org.

ACADEMY OF BANKING STUDIES JOURNAL

CONTENTS OF VOLUME 9, NUMBER 1

ACADEMY OF BANKING STUDIES JOURNAL

CONTENTS OF VOLUME 9, NUMBER 2

LETTER FROM THE EDITORS

The *Academy of Banking 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 *ABSJ* is a principal vehicle for achieving the objectives of the organization.

The *ABSJ* has undergone a name change effective with this issue. It was formerly known as the *Journal of Commercial Banking and Finance*. We have changed its name to more closely match its editorial mission, which is to publish theoretical and empirical manuscripts which further the disciplines of banking and institutional finance. The name change resulted in the issuance of a new ISSN, 1939-2230. The former ISSN was 1544-0028. The journal continues to follow its 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 *ABSJ*, 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.

Kurt Jesswein
Sam Houston State University

This is a combined edition containing both Volume 9, Number 1, and Volume 9, Number 2

Articles for Volume 9, Number 1

A DISCRETE TIME MARKOV CHAIN MODEL FOR PREDICTING THE DURATION OF A RETAIL MORTGAGE IN THE NON-DEFAULT STATES

Morsheda Hassan, Wiley College
Chang Liu, SouthWest University of Finance and Economics, China
Raja Nassar, Louisiana Tech University

ABSTRACT

In this study, a discrete time Markov chain model is developed for modeling the duration of retail loans with prepayment, past due, and default states. Prepayment and past due states describe the payment status of a loan. The default state is defined as charge-off on the loan due to bankruptcy, death, or other causes. .A bank could use this model to predict its contingent assets status based on the probability and duration of being in the non-default states. This is vital for effective management of a Bank's assets.

INTRODUCTION

Duration of a retail mortgage, which could provide a snapshot of a portfolio' health status, is defined as the expected number of days before a loan is charged off. Te terms or number of years that are usually specified by the mortgage contract could not be considered as an effective proxy for the expected duration defined above. Ths is because of the extension risk, which is defined as the extension of the mortgage terms because of the past due status of the loan, and the contraction risk, which is defined as the shrinkage of the mortgage terms because of the early payment, or prepayment as most mortgage contracts do not include the prepayment penalty. As a result, an accurate prediction of duration, taking both extension risk and contraction risk into account, could be a major factor in credit risk management.

In this study, the authors propose a discrete Markov chain model to predict the expected duration of the loan in the non-default states. This model could be used to predict the timing of a given credit event for portfolios with different vintages, and thus provide a dynamic comparison between portfolios. For example, a 20-year semi-paid portfolio with expected default at the 21st year is definitely better than a 30-year annually paid portfolio with expected default at the 29th year.

The structure of this study is organized as follows: Section one provides a review of salient studies using Markov chain models to perform credit analysis. Section two gives a theoretical

derivation of the portfolio economic assets model based on a Markov chain. Section three provides an empirical application of the model, and section four gives the conclusion and future studies.

LITERATURE REVIEW

There are many quantitative methods in credit asset management. As summarized by White (1993), Markov decision models have been frequently used in 18 areas, including (1) Finance and Investment, (2) Insurance, and (3) Credit Analysis. Of the 98 papers discussed by White, 9 papers relate to finance and investment, 2 to insurance, and 2 to credit analysis. This survey is by no means comprehensive, but it reveals the fact that Markov chains have been used extensively t analyze real world data.

General concepts of Markov processes are presented in Ross (1996). Let π_{ij} be the steady state probability or limiting probability of being in state i and adapting policy j, $\pi_{ij} = lim_{n\to\infty} P_{ij}(n)$, where X_n, n=1,2,3…represent the state of a Markov chain at the *nth* transition.. As such, the expected benefit is given as

$$\Sigma_i \Sigma_j \pi_{ij}[R(i, j)-C(i, j)] \quad (1)$$

where, R(i, j), C(i, j)are defined as the reward function and cost function for being in state i and adopting policy j, respectively. Also, dynamic programming could be used to find an optimal policy j to maximize the expected benefit. To this end, one may maximize

$$\Sigma_i \Sigma_j \pi_{ij}[R(i, j)-C(i, j)]$$
$$\text{Subject to } \pi_{ij} \geq 0, \text{ and } \Sigma_i \Sigma_j \pi ij=1 \quad (2)$$

Consumer credit analysis is used to analyze account receivable, as triggered by credit sales. The model, based on the transition probability between different states, is primarily used by a company to adjust its credit sale and collection policy. Absorbing states could be reached either by collection or bad debt, both of which lead to a decline in the portfolio size.

On the other hand, by defining a past-due period as a different transient state, and default as an absorbing state, Markov models are used to analyze the characteristics of a loan portfolio, namely the estimated duration before an individual default, prediction of economic portfolio balance, and health index. The primary purpose of this study is to develop this type of model for banks and other commercial lending institutes in order to analyze the nature of their products.

Markov Models for Consumer Credit Analysis

Cyert, Davison and Thompson (1962) developed a finite stationary Markov chain model to predict uncollectible amounts (receivables) in each of the past due category. This classic model is referred to as the CDT model. The states of the chain (S_j, $j=0,1,2\ldots,J$) were defined as normal payment, past due, and bad-debt states. The probability P_{ij} of a dollar in state S_i at time t transiting to state S_j at time t + 1 is given as

$$P_{ij}=B_{ij}/\Sigma_{m=0\ to\ J}B_{im} \qquad (3)$$

where B_{ij} is the amount in state S_j at time t+1 which came from state S_j in the previous period. $S_t(j)=S_0(j)Q^t$ is the vector whose *jth* component is the amount outstanding for the *jth* past due category at the beginning of the *tth* period for $t=1,2,\ldots$, Here, Q is a sub-matrix, in the transition probability matrix $P_{ij}=[I\ O;\ R\ Q]$, which includes transition probabilities among the set of transient states.

Criticizing the appropriateness of the stationary Markov chain model by Cyert, Davidson and Thompson, (1962), Frydman (1985) applied a mover-stayer Model as an alternative. They defined the j step transition matrix of this model as $P(0,j)=SI+(I-S)M^j$, where $M=\{m_{ik}\}$ is a transition probability matrix for "movers" from i to k, and $S = diag(s_1,s_2,\ldots,s_w)$ represents the probability of "stayers" in state i. The maximum likelihood estimator for m_{ik} is given as $m_{ik}=(n_{ii}\text{-}Jn_i)(n_i\text{*-}Jn_i)$, where n_i is the number of observations that stay in state i during the period. They concluded that the mover-stayer model is better for empirical analysis than the stationary Markov chain model

The model of Cyert et al.(1962) was also challenged by Corcoran (1978). Corcoran claimed that the representiveness of the transition probability could be affected by the fact of "dominancy of large accounts". Therefore, he suggested grouping the accounts according to their size, and a transition matrix for each group was provided by an exponentially smoothed matrix: $A_j=0.8T_j+(1\text{-}0.8)A_{j\text{-}1}$, where A_j is an exponentially smoothed matrix for month j and T_j is the transition matrix for month j.

Markov Models for Loan Analysis

According to Thompson (1965), one of two important related tests for a bank's credit asset from the lender's point of view is the possibility of the loan getting into trouble, which means the probability of being in a past-due or even charged-off state. Another test is the extent of loss in the case of being in trouble. This could mean two things: (1) the recovery from collateral in the case of being charged off, or (2) the ability for an individual to bring himself back on track. Also, in the same paper, Thompson provided evidence supporting his claim that the business cycle and the macroeconomic situation are probably the most significant factors affecting change in bank credit.

Howard and Matheson (1972) implemented a Markov model which could be useful in forming optimum buying and selling strategies for a commodity market. They justified the model by incorporating a risk-sensitivity function. A positive or negative risk coefficient was assigned to the function based on whether the bank management is risk aversive or risk preferring, respectively. The exponential function representing the overall risk preference is given by: $u(v)=-(sgnr)e^{-rv}$, where r is the risk aversion coefficient, and *sgnr* denotes the sign of r. Then, the following iterative scheme was provided to reach the maximum benefits through an optimal policy:

$$F(k)=\Sigma_{j=1 \text{ to } N} p_{ij} e^{-r\{r_{ij} + -(1/r)ln[-(sgnr)u_i(v\ k)]\}} \quad (4a)$$

$$\text{Choose policy } k \text{ to maximize } V_i^k=-(1/r)ln[\Sigma_{j=1 \text{ to } N} p^k_{ij} F(k)] \quad (4b)$$

where, p_{ij} is the Markov transition probability, and $F(k)=e^{-r\{r_{ij} + -(1/r)ln[-(sgnr)u_i(v\ k)]\}}$. For an initial risk preference function, $u(v)$, given by credit policy $k \in \{\kappa\}$, where κ is the set of all feasible credit policy packages available for the bank to select from, the bank risk manager can use Eq.(6a) to get $F(k)$, and substitute it into Eq.(6b) to find V_i^k. This iteration scheme will generate the best credit policy.

By taking economic factors into account, Richard (1983) used a finite Markov chain model to analyze a firm's market value if the firm follows an optimal policy in *state(x,y)* at time *t*, where *x* is the condition of the firm, and *y* is the condition of the overall economy. He assumed that the changes in state are governed by a stationary transition function. For instance, if the state is *y(t*-1*)* at time *t*-1, then it will be *y(t)* at time *t* with probability $\pi[y(t)\ y(t\text{-}1)]$. However, to calculate V_i^k, he used dynamic programming because direct computation could be very time-consuming.

Jarrow (1997) applied a continuous and a discrete time Markov chains to describe the default behavior of zero-coupon bonds within a time interval $\eta t(0 \le t \le)$. Furthermore, the default state was defined as an absorbing state. Again, the purpose was to price the bond based on analysis of credit risk spread. Similar approaches have been adopted by Liebman (1972) and Zipkin (1993). Glennon and Nigro (2005) used the survival analysis approach to measure the default risk of a small business. They adopted the Cox Proportional Hazard model. By using a discrete-time hazard procedure, they found that the default risk peaked in the second year after initiation, increased during the medium-maturity season, and declined thereafter.

Guedon (2003) studied the algorithms for estimation of hidden semi-Markov chains for nonstationary discrete time series data in the framework of expectation-maximization (EM) algorithm. He emphasized determining the appropriate number of states. Similar models have been implemented by Shun and Hisashi (2002) for signal processing and by Pieczynski (2005) for unsupervised segmentation of the hidden non stationary semi-Markov chains.

Instead of solving the problem analytically, Sandmann (2005) proposed a variance reduction technique for efficient simulation by importance sampling. He tested the optimal forms of importance sampling for estimating state probabilities in discrete-time Markov chains.

MODELS

In the Markov chain model considered in this study, let S_j be a state of past due, corresponding to the days of past due. The loan normally requires monthly payment. If a loan is 30 days past due, denote it by S_1. State S_2 refers to 60 days past due. According to the Basel accord II, (Basel Committee on Banking Supervision, 1997), the definition of default is more than 90 days past due, which is represented by S_3. However, there have been cases where the obligations on a loan, which have already been more than 90 days past due, has been paid off. As a result, the definition of default is modified to be the state of default that is triggered by a permanent force, such as death or an application of chapter 7 or chapter 13 bankruptcy protections. Let S_i be the default state contributed by these permanent events and let S_{-i} be the state of a prepaid period defined as $S_{-j}=(X_i-Y_i)$, where X_i is the actual payment at month i and Y_i is the scheduled payment at month i. One can see that state S_{-j} is defined as the extra payment over the scheduled payment, which measures how many future monthly payments have been made as a current onetime payment. It is not a precise measurement method, compared with the tools introduced by other papers in the literature, but it fits best in the context of this model. Classification of the states of the Markov chain are given below.

Definitions of the different states of the Markov chain.

We follow the approach in Kemeny and Snell (1983) to evaluate the expected number of times, $e_{ij}(t)$, the process remains in the non-absorbing state S_i once this state is entered, including the entering step. If the expected number of times the process stays in the non-absorbing state S_i, which is defined as a non-default state, is n_i, then the expected duration of stay in S_i is n_i*30, where 30 days is the step size of the transition probability matrix. As such, the expected total number of days the process is in the non-default states is $\Sigma_{i=-3\ to\ 3}\ n_i*30$. To facilitate the computation, we define V as the overall transition matrix, including prepayment, past due, and default states. The one-step transition probability matrix is given in Figure 1.

Here, v_{ij} refers to the probability of transition from S_i to S_j, μ_{ik} refers to the probability of transition from S_i to R_k. Furthermore, because R_k is an absorbing state, the transition matrix for these states is an Identity matrix and μ_{ki} =0. In the V matrix, there are 7 transient states (S_i's)in which each state could be reached from any other state, and 4 absorbing states (R_k's) each of which can be reached from any transient state. Once an absorbing state is entered the process remains in that state and cannot exit. The V matrix could be rearranged into block matrices as shown in Figure 2:

Figure 1 One-step transition probabilities.

$$V = \begin{array}{c} \\ R_1 \\ R_2 \\ R_3 \\ R_4 \\ S_{-3} \\ S_{-2} \\ S_{-1} \\ S_0 \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{ccccccccccc} R_1 & R_2 & R_3 & R_4 & S_{-3} & S_{-2} & S_{-1} & S_0 & S_1 & S_2 & S_3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mu_{-3,1} & \mu_{-3,2} & \mu_{-3,3} & \mu_{-3,4} & \nu_{-3,-3} & \nu_{-3,-2} & \nu_{-3,-1} & \nu_{-3,0} & \nu_{-3,1} & 0 & 0 \\ \mu_{-2,1} & \mu_{-2,2} & \mu_{-2,3} & \mu_{-2,4} & \nu_{-2,-3} & \nu_{-2,-2} & \nu_{-2,-1} & \nu_{-2,0} & \nu_{-2,1} & 0 & 0 \\ \mu_{-1,1} & \mu_{-1,2} & \mu_{-1,3} & \mu_{-1,4} & \nu_{-1,-3} & \nu_{-1,-2} & \nu_{-1,-1} & \nu_{-1,0} & \nu_{-1,1} & 0 & 0 \\ \mu_{0,1} & \mu_{0,2} & \mu_{0,3} & \mu_{0,4} & \nu_{0,-3} & \nu_{0,-2} & \nu_{0,-1} & \nu_{0,0} & \nu_{0,1} & 0 & 0 \\ \mu_{1,1} & \mu_{1,2} & \mu_{1,3} & \mu_{1,4} & \nu_{1,-3} & \nu_{1,-2} & \nu_{1,-1} & \nu_{1,0} & \nu_{1,1} & \nu_{1,2} & 0 \\ \mu_{2,1} & \mu_{2,2} & \mu_{2,3} & \mu_{2,4} & \nu_{2,-3} & \nu_{2,-2} & \nu_{2,-1} & \nu_{2,0} & \nu_{2,1} & \nu_{2,2} & \nu_{2,3} \\ \mu_{3,1} & \mu_{3,2} & \mu_{3,3} & \mu_{3,4} & \nu_{3,-3} & \nu_{3,-2} & \nu_{3,-1} & \nu_{3,0} & \nu_{3,1} & \nu_{3,2} & \nu_{3,3} \end{bmatrix} \end{array}$$

Figure 2 One-step transition probability matrix and its block matrices.

$$V = \left[\begin{array}{c|c} I & O \\ \hline R & Q \end{array}\right] \tag{5}$$

$$I_{4\times4} = \begin{array}{c} \\ R_1 \\ R_2 \\ R_3 \\ R_4 \end{array} \begin{array}{c} \begin{array}{cccc} R_1 & R_2 & R_3 & R_4 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}, \quad O_{4\times7} = \begin{array}{c} \\ R_1 \\ R_2 \\ R_3 \\ R_4 \end{array} \begin{array}{c} \begin{array}{ccccccc} S_{-3} & S_{-2} & S_{-1} & S_0 & S_1 & S_2 & S_3 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

$$R_{7\times4} = \begin{array}{c} \\ S_{-3} \\ S_{-2} \\ S_{-1} \\ S_0 \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{cccc} R_1 & R_2 & R_3 & R_4 \end{array} \\ \begin{bmatrix} \mu_{-3,1} & \mu_{-3,2} & \mu_{-3,3} & \mu_{-3,4} \\ \mu_{-2,1} & \mu_{-2,2} & \mu_{-2,3} & \mu_{-2,4} \\ \mu_{-1,1} & \mu_{-1,2} & \mu_{-1,3} & \mu_{-1,4} \\ \mu_{0,1} & \mu_{0,2} & \mu_{0,3} & \mu_{0,4} \\ \mu_{1,1} & \mu_{1,2} & \mu_{1,3} & \mu_{1,4} \\ \mu_{2,1} & \mu_{2,2} & \mu_{2,3} & \mu_{2,4} \\ \mu_{3,1} & \mu_{3,2} & \mu_{3,3} & \mu_{3,4} \end{bmatrix} \end{array}, \quad Q_{7\times7} = \begin{array}{c} \\ S_{-3} \\ S_{-2} \\ S_{-1} \\ S_0 \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{ccccccc} S_{-3} & S_{-2} & S_{-1} & S_0 & S_1 & S_2 & S_3 \end{array} \\ \begin{bmatrix} \nu_{-3,-3} & \nu_{-3,-2} & \nu_{-3,-1} & \nu_{-3,0} & \nu_{-3,1} & 0 & 0 \\ \nu_{-2,-3} & \nu_{-2,-2} & \nu_{-2,-1} & \nu_{-2,0} & \nu_{-2,1} & 0 & 0 \\ \nu_{-1,-3} & \nu_{-1,-2} & \nu_{-1,-1} & \nu_{-1,0} & \nu_{-1,1} & 0 & 0 \\ \nu_{0,-3} & \nu_{0,-2} & \nu_{0,-1} & \nu_{0,0} & \nu_{0,1} & 0 & 0 \\ \nu_{1,-3} & \nu_{1,-2} & \nu_{1,-1} & \nu_{1,0} & \nu_{1,1} & \nu_{1,2} & 0 \\ \nu_{2,-3} & \nu_{2,-2} & \nu_{2,-1} & \nu_{2,0} & \nu_{2,1} & \nu_{2,2} & \nu_{2,3} \\ \nu_{3,-3} & \nu_{3,-2} & \nu_{3,-1} & \nu_{3,0} & \nu_{3,1} & \nu_{3,2} & \nu_{3,3} \end{bmatrix} \end{array}$$

. . . . . .

Thus, several interesting results could be reached. First, the expected number and variance of the time the process stays in a non-absorbing state S_i, n_i , before leaving S_i, are given by:

$$E(n_i)=N\zeta$$
$$Var(n_i)=(2N-I)\ n_i - sq(n_i) \quad (6)$$

where,

$N=(I-Q)^{-1}$, I is an 7*7 identity matrix.
$sq(n_i)=[n^2_{-3}\ \ n^2_{-2}\ \ n^2_{-1}\ \ n^2_{0}\ \ n^2_{1}\ \ n^2_{2}\ \ n^2_{3}]$
ζ=[1 1 1 1 1 1 1]'

In fact, $E_i(n_j)$, the expected number of steps needed to transit from state i to state j, is equal to N_{ij}.

Let
$$d_{ij}= \begin{cases} 1, \text{ if } i=j \\ 0, \text{ if } i \neq j \end{cases} \quad (7)$$

Thus, it is seen that: $E_i(n_j)= d_{ij} + \Sigma_{\in \square I} p_{ik}E_k(n_j)$, which, in matrix form, is:

$$E=\{E_i(n_j)\}=I+QE_i(n_j)$$
$$E=\{E_i(n_j)\}=(I-Q)^{-1}=N \quad (8)$$

Where $s \square T$. By summing over i from -3,-2,-1,0,1,2,3, one obtains the total expected number of steps the process is in a non-absorbing state S_i before going to a default state R_k. Then, the expected total days of stay in the non-default states should be:

$$\Sigma_{i=-3\ to\ 3} E_i(n_j)*30 \quad (9)$$

APPLICATION

Data were provided by a local bank in Ohio, operating in Ohio, Michigan, Kentucky, and Indiana. By using its monthly paid retail mortgage loan for 16 consecutive months, from April 2005 to September 2006, one can demonstrate the applicability of the model.

As shown in Figure 2, the intensity matrix V, can be divided to 4 sub-matrices. These are an identity matrix I_{4*4}, a zero matrix O_{4*7}, a R_{7*4} matrix which refers to the transitions from transient to absorbing states, and the Q_{7*7} matrix which denotes transitions within the transient states. By the definition of an absorbing state, it is seen that the intensity sub-matrix within the absorbing states is an identity matrix because once entered into an absorbing state, the loan will stay there for an infinite period of time. By the same reasoning, the zero matrix O_{4*7} refers to the fact that there is no transition from any absorbing state to any transient state.

On the other hand, elements of the transient Q_{7*7} matrix and the ergodic matrix R_{7*4} can be estimated as

$$Q_{ij}=(\Sigma_{t=1 \text{ to } 15}\, q_{ijt})/(\Sigma_{t=1 \text{ to } 15}\Sigma_{j=-3 \text{ to } 3} q_{ijt});$$
$$R_{ik}=(\Sigma_{t=1 \text{ to } 15}\, r_{ikt})/(\Sigma_{t=1 \text{ to } 15}\Sigma_{k=1 \text{ to } 4} q_{ikt});$$
$$i, j = -3,-2,-1,0,1,2,3;\ k=1,2,3,4, \qquad (10)$$

where t refers to the consecutive time intervals $\{m,m+1\}$,m=1,2,…,16 months .

For example, the number of transitions between period 1 and period 2 (t =1→2) are given in Figure 3.

Figure 3. Two consecutive month transition matrices.

$$q_1 := \begin{pmatrix} 10 & 3 & 4 & 3 & 0 & 0 & 0 \\ 1 & 9 & 8 & 1 & 0 & 0 & 0 \\ 1 & 1 & 51 & 21 & 1 & 0 & 0 \\ 3 & 9 & 32 & 722 & 12 & 0 & 0 \\ 0 & 1 & 2 & 6 & 8 & 2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 3 \end{pmatrix} \qquad r_1 := \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 12 & 0 & 15 \\ 0 & 0 & 1 & 10 \\ 0 & 4 & 1 & 0 \end{pmatrix}$$

The numbers in the matrices represent the number of transitions from state i to state $j(q_{ij1})$, i, j =-3,-2,-1,0,1,2,3, or from state i to state $k(r_{ik1})$, k=1,2,3,4,between period 1 and period 2. The transition probabilities as estimated from (10) are given in Figure 4.

Figure 4 Discrete transition probability matrices

$$Q := \begin{pmatrix} 0.575949 & 0.09962 & 0.164557 & 0.132911 & 0 & 0 & 0 \\ 0.036496 & 0.61365 & 0.156277 & 0.182482 & 0.000562 & 0 & 0 \\ 0.027559 & 0.034121 & 0.677428 & 0.239895 & 0.001452 & 0 & 0 \\ 0.004084 & 0.009734 & 0.049125 & 0.89205 & 0.010912 & 0 & 0 \\ 0.009094 & 0.008955 & 0.01403 & 0.207812 & 0.206307 & 0.055224 & 0 \\ 0 & 0.003157 & 0.023684 & 0.034737 & 0.085263 & 0.174737 & 0.318421 \\ 0 & 0 & 0.002105 & 0.0632 & 0.09526 & 0.151053 & 0.293157 \end{pmatrix}$$

$$R := \begin{pmatrix} 0.019452 & 0 & 0 & 0 \\ 0.010451 & 0 & 0 & 0 \\ 0.010937 & 0 & 0 & 0 \\ 0.029017 & 0 & 0 & 0 \\ 0.132985 & 0.152321 & 0 & 0.209015 \\ 0 & 0.090526 & 0.145632 & 0.122053 \\ 0 & 0.105206 & 0.150526 & 0.130526 \end{pmatrix}$$

$$I := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad O := \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Then, according to Kemeny and Snell (1983), the overall intensity matrix *V*, composed of R_{7*4}, Q_{7*7}, I_{4*4}, $O_{4*7}$is given in Figure 5:

Figure 5 Discrete transition probability matrix *V*.

$$V = \begin{array}{c|cccc|ccccccc} & R_1 & R_2 & R_3 & R_4 & S_{-3} & S_{-2} & S_{-1} & S_0 & S_1 & S_2 & S_3 \\ \hline R_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ R_2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ R_3 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ R_4 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline S_{-3} & 0.019452 & 0 & 0 & 0 & 0.575949 & 0.09962 & 0.164557 & 0.132911 & 0 & 0 & 0 \\ S_{-2} & 0.010451 & 0 & 0 & 0 & 0.036496 & 0.61365 & 0.156277 & 0.182482 & 0 & 0 & 0 \\ S_{-1} & 0.010937 & 0 & 0 & 0 & 0.027559 & 0.034121 & 0.677428 & 0.239895 & 0.001452 & 0 & 0 \\ S_0 & 0.002917 & 0 & 0 & 0 & 0.004084 & 0.009734 & 0.049125 & 0.892050 & 0.010912 & 0 & 0 \\ S_1 & 0.132985 & 0.152321 & 0 & 0.209015 & 0.009094 & 0.008955 & 0.01403 & 0.207812 & 0.206307 & 0.055224 & 0 \\ S_2 & 0 & 0.090526 & 0.145632 & 0.122053 & 0 & 0.031579 & 0.023684 & 0.034737 & 0.085263 & 0.174737 & 0.318421 \\ S_3 & 0 & 0.010526 & 0.010526 & 0.130526 & 0 & 0 & 0.002105 & 0.063200 & 0.095260 & 0.151053 & 0.293157 \end{array}$$

Thus, by the above reasoning, the expected number of steps required to transition from transient state *i*, *i*=-3,-2,-1,0,1,2,3 to an absorbing state *k, k*=1,2,3,4 is given by *E*:

$$E=(I-Q)^{-1}=N\zeta=[29.38866\ 30.16227\ 29.24377\ 27.42631\ 10.12164\ 6.99529\ 6.81304]' \quad (11)$$

In this case, one can see that it takes 29.38866 or approximately 30 steps for a loan initially in state -3 to leave the transient states for any absorbing state. In other words, since the step is 1 month, a loan more than 3 months prepaid (state 3) could become sold or defaulted in approximately 30 months or 2.5 years, while a loan with 3 months past due could reach the same destiny in approximately 7 months.

Also, let b_{ik} be the probability that the process transits from transient state *i, i* =-3,-2,-1,0,1,2,3 to absorbing state *k, k*=1,2,3,4:

$$\{b_{ij}\}=B=(I-Q)^{-1}R=NR=\begin{bmatrix} 0.72337 & 0.04473 & 0.00429 & 0.06121 \\ 0.72953 & 0.04692 & 0.0045 & 0.0642 \\ 0.71614 & 0.04673 & 0.00448 & 0.06395 \\ 0.72809 & 0.04968 & 0.00476 & 0.06798 \\ 0.39812 & 0.22249 & 0.02133 & 0.30446 \\ 0.1545 & 0.22566 & 0.28605 & 0.2984 \\ 0.1539 & 0.23163 & 0.2774 & 0.29573 \end{bmatrix} \quad (12)$$

An element of *B,* b_{ik}, represents the probability of transiting from transient state *i* to absorbing state *k.* For example, b_{32}=0.23163 means that the probability of transition from the 3-month past due state to the absorbing state (foreclosure) is 0.23163.

CONCLUSION

The above model describes the expected retail mortgage loan's behavior. Furthermore, this model provides useful information to quantify the risks encountered by the bank's management. By using this model, the management can obtain a clear picture of its retail loans. For example, from Eq (11), we know approximately how long the loan belonging to a certain state could take to enter an absorbing state. Thus, a corresponding rescue action could be deployed to encounter each situation.

Nevertheless, the discrete time Markov model is by no means the only tool that could be deployed by the bank management. In fact, the above model uses only the occurrence frequencies of each state and did not consider the loan assets which are important for risk management in the banking industry. In future work, we will extend the present study to include loan assets.

REFERENCES

Basel Committee on Banking Supervision. *International Convergence of Capital Measurement and Capital Standards*, CH-4002 Basel 1997.

Corcoran, Wayne (1978). The Use of Exponentially-Smoothed Transition Matrices to Improve Forecasting of Cash Flows from Accounts Receivable. *Management Science* 1978; 23: 45-54.

Cyert M, Davidson H, and Thompson G (1962). Estimation of the Allowance for Doubtful Accounts by Markov Chains. *Management Science* 1962; 8: 3-19.

Frydman, Halina, Jarl Kallberg, and Duen-Li Kao (1985) Testing the Adequacy of Markov Chain and Mover-Stayer Models as Representations of Credit Behavior. *Operations Research*, 12: 3-16.

Glennon, Dennis and Peter Nigro (2005). Measuring the Default Risk of Small Business Loans: A Survival Analysis Approach . *Journal of Money, Credit, and Banking* 2005; 37:109-125.

Grinold, Richard.(1983), Market Value Maximization and Markov Dynamic Programming, *Management Science*, 1983;14:123-145

Guedon, Yann (2003. Estimating hidden semi-Markov chains from discrete sequences. *Journal of Computational & Graphical Statistics*, Sep. 03

Howard, Ronald and James Matheson. (1972). Risk-Sensitive Markov Decision Processes. *Management Science*; 18: 7-14.

Jarrow, Robert, David Lando, and Stuart Turnbull. (1997). A Markov Model for the Term Structure of Credit Risk Spreads. *The Review of Financial Studies*; 10: 231-239

Kemeny, John and Laurie Snell (1983). *Finite Markov Chains,* D. Van Nostrand Company.

Liebman, Leon (1972). A Markov Decision Model for Selecting Optimal Credit Control Policie. *Management Science*; 12: 261-289.

Pieczynski, Wojciech (2005), Modeling Nonstationary Hidden Semi-Markov Chains with Triplet Markov Chains and Theory of Evidence. IEEE Workshop on Statistical Signal Processing (SSP2005), Bordeaux, France, July 2005

Ross, Sheldon (1996). *Introduction to Probability Models,* Academic Press.

Sandman, Werner (2005), On optimal importance sampling for discrete-time Markov chains, *Quantitative Evaluation of Systems*, Volume 19 Issue 22

Thompson, Donald (1956), Changes in Quality of Bank Credit. *The Journal of Finance* 1956; 11: 2-45.

White, Joyce (1993). A Survey of Applications of Markov Decision Process. *The Journal of the Operational Research Society*; 44: 11-20.

Yu, Shun-Zheng and Hisashi Kobayashi (2002). A hidden semi-Markov model with missing data and multiple observation sequences for mobility tracking, *Signal Processing,* Volume 83, Issue 2

Zipkin, Paul. (1993). Mortgages and Markov Chains: A Simplified Evaluation Model. *Management Science* , 39: 324-332.

AN ANALYSIS OF PRODUCTIVITY CHANGE: ARE UAE BANKS OPERATING EFFICIENTLY WHEN COMPARED TO GCC BANKS?

Hossein Attar Kashani, University of Dubai
Lamia Abdelaziz Obay, University of Abu Dhabi

ABSTRACT

This study, for the first time, takes on the issue of productivity changes of UAE banks in comparison to those in the rest of the Gulf Cooperation Council (GCC) countries over the period 2000-2005. Given the growing importance of Islamic banking, this study also examines the difference in efficiency between Islamic and Conventional banks. Based on the non-parametric approaches of data envelopment analysis, DEA, and Malmquist productivity index, MPI, the results show that the banks in the GCC countries show relatively similar levels of efficiency. While Kuwait and Qatar had higher efficiency scores than the UAE and the UAE higher than the rest of the remaining GCC countries, these were not found to be statistically significant. Commercial Bank of Abu Dhabi and National Bank of Dubai appeared constantly on the best practice frontier. The statistical tests also show no significant differences between the performance of the Islamic banks and their conventional counterparts. Over time, UAE banks were able to show gains in efficiency (4%) when other GCC banks were actually recording losses of efficiency of the same magnitude. Both Islamic and Conventional banks in the UAE recorded gains in technical and pure efficiency and losses in scale efficiency.

INTRODUCTION

Efficiency analysis of commercial banks is particularly needed in a region dominated by bank-intermediated finance. A well developed banking sector was found to bear a positive relationship to economic growth (Levine and Zervos, 1998). It was further suggested that the legal environment within which banks operate can significantly affect economic growth through its effect on bank behavior (Levine, 1998).

The economies of the Gulf Cooperation Council (GCC) have gone through periods of peaks and troughs, mainly because of their dependence on oil revenues and the volatility of oil prices. After a period of economic slump due to declining oil prices during most of the 1990s, GCC countries have been witnessing breathtaking growth thanks to skyrocketing oil prices since 2000. The surge in oil revenues has led to a lifting of the region's economy and the accumulation of large

amount of liquidity. This has stimulated an unprecedented investment boom, especially in the property market, and increasing demand for bank services. The question becomes whether banks are well poised to meet the challenges and take advantage of the new opportunities presented to them.

Banks in the region are generally small in size, preventing them from participating fully in the financing of energy-related and other local infrastructure projects. They have long benefited from the protectionist policies of the local governments. Entry and branching restrictions have limited the ability of foreign banks to capture a material share of the local loan and deposit market, even when outnumbering the local ones.

Bank regulators, however, have expressed intentions to, or starts of, changes, which, once completely implemented, would confront the banks – specially the local ones – with serious challenges. To start with, all the GCC member countries are signatories of the WTO agreement, which should result in further liberalization of their economies, in general, and the financial sector in particular. Indeed, most of these six countries have already revised, or are in the process of revising, their "company and investment laws" to allow for higher level of foreign ownership of banks and presence of foreign investors in the local stock markets. It is these changes that made Kuwait, Oman and Saudi Arabia extend new licenses to foreign banks in 2004. The UAE recently indicated its intention to adopt a reciprocal treatment to foreign banking presence when issuing new banking licenses.

Further, the GCC countries are envisioning a monetary union by the year 2010. This not only assumes the removal of all barriers towards the flow of capital between the member countries. It may also lead, through mergers and acquisition, to a wave of consolidation that would unveil hidden inefficiencies that were only made viable because of government protectionist policies.
One salient feature common to all GCC countries is the emergence of Islamic banking and finance. It all started with the establishment of Dubai Islamic Bank which was able to bring about a marriage of faith and finance that many thought could not coexist in modern times. While the Islamic finance sector has been growing at a faster rate than its conventional counterpart, it still is far from capturing a sizable market share of the banking industry. During 2001-2005, the UAE Islamic Banks outperformed their GCC counterparts in terms of growth of their asset base (20.5%), loan portfolio (21.6%), and deposit mix (21.2%) compared to GCC's growth rates of 7%, 6% and 4% respectively. Islamic banks in Bahrain lead in terms of market share: viz., 38% of the country's banking sector real assets, 28% of real loans, and 36% of real deposits in 2005 compared with 12.4%, 16.1% and 13.6% respectively in the UAE.

With these scenarios, the efficiency and/or over-the-time efficiency gain of each bank in the system would play a major role in their competitiveness and survival as an independent entity, for the less efficient firms have traditionally been the prime target of well-functioning competitors. On a cross-country basis, on the other hand, the possibility of a monetary union between the member countries would suggest homogenous banking system in each and every GCC countries. Without this homogeneity cross-country mergers and acquisitions are possibilities that could not be ignored.

Once again the efficiency of banks and their over-the-time efficiency gain play major role in provoking a take-over initiation.

This study is, therefore, important for bankers as a higher efficiency means higher profit and increased chance of survival in an increasingly deregulated and competitive market. Higher efficiency could also lead to a higher customer satisfaction as efficient banks are better positioned to offer quality and new services at competitive prices. The study is also important for the policy makers. An awareness of efficiency features is important to help them formulate policies that affect the banking industry as a whole and the local banks in particular.

This study aims to examine the overall cost efficiency of the banks in all six GCC member countries of the GCC over the period of 2000-2005 and the comparative efficiency between Islamic banks (IBs) and conventional banks (CBs). We measure the efficiency of each bank and the average efficiency of the banks in each country and for each specialization (IBs vs CBs) in each year of the operation within the time period in question. We proceed by investigating the efficiency gain/loss of each bank during the mentioned period to shed some lights on overall performance of the banking industry in the GCC countries.

We also test whether there has been a significant improvement in their efficiency over time. Given the protectionist environment in which they operate, the high level of government ownership in the sector, the low level of financial deepening of the GCC economies and the most recent oil bonanza, banks may have had little incentive to strive towards improving their productive efficiency, and accordingly, achieve little efficiency gains.

The remainder of this study is organized as follows. The next section looks at some of the existing works on the issue. This will be followed by discussing the methodologies used in this study. We will see how data envelopment analysis can be used to measure the relative efficiency of the banks and how we can apply the Malmquist productivity index (MPI) technique to break down the efficiency changes in various components and how it could be used to measure the efficiency changes of the banks over time. To measure the efficiency of the banks, we need the input and output data. Section four discusses the variables we use in this study. We then proceed to introduce the results. The study comes to its end with a summary and some concluding remarks.

LITERATURE

The literature focusing on the efficiency of the financial sectors of various countries, in general, and the banking sector, in particular, is vast. To mention only a few of the literature during the last 10 years we can name: Sufian and Abdul Majid (2007) discusses the relationship between X-efficiency and share prices in the Singaporean banking sector; Sufian (2007) evaluates the efficiency of domestic and foreign Islamic banks active in Malaysia; Barros and Garcia (2006) use DEA to evaluate the performance of Portuguese pension funds from 1994 to 2003; Lozano-Vivas and Pastor (2006) relate macro-economic efficiency of fifteen OECD countries over a period of eighteen years to the financial efficiency of the countries; Grigorian and Manole (2006) use DEA

to estimate indicators of commercial bank efficiency to bank-level data from a wide range of transition countries; Brown and Skully (2006) evaluate the cost efficiency of banks in the Asia-Pacific region and test whether the operating performance of banks in poorer economies improves with the inclusion of environmental proxies; Kirkwood and Nahm (2006) investigate the relationship between the Australian banks' efficiency to their stock returns; Lo and Lu (2006) discuss the profitability and marketability benchmark of financial holding companies in Taiwan; Samjeev (2006) evaluates the efficiency of the public sector banks in India to investigate whether there exists any relationship between the efficiency and size of the banks; Camanho and Dyson (2005) use DEA to measure the cost efficiency of a British bank branches in phase 1 to be applied in the analysis of branch network and the production and value-added approaches to have a more comprehensive assessment of bank branch efficiency; Weill (2004) investigates the consistency of efficiency frontier models on some European (France, Germany, Italy, Spain, and Switzerland) banking samples; Krishnasamy, Hanuum Ridzwa, and Perumal (2003) apply Malmquist Productivity Index to discuss the efficiency of the Malaysian banks' post-merger productivity; Sathye (2002) measures the productivity changes in Australian banking sector using DEA and Malmquist Productivity Index; Athanassopoulos and Giokas (2000) use DEA to discuss the efficiency of the banking sector in Greece; Chen and Yeh (2000) measure the bank efficiency and productivity changes in Taiwan banking sector and investigate the impact of ownership on the resulted efficiency scores; Zenios et. al. (1999) use DEA to develop a benchmark on the relative efficiency of the Cyprus banks branches, provide guidelines for improvement to management, and isolate the effects of the environment on branch efficiency; Camanho and Dyson (1999) too use DEA to assess the performance of Portuguese bank branches to complement the profitability measure used by the bank; Ayadi, Adebayo, and Omolehinwa (1998) measure the bank performance of Nigerian banks and conclude that the seeming inefficiency is attributed to the banks' poor management;

Closer to the region under this study, Rao (2005) looks at 35 banks operating in the UAE for the years 1998 and 2000, and concludes that these banks suffer substantial cost, X- (managerial), and scope inefficiency. Saif and Yaseen (2005) look at Scope and Scale efficiency of 100 banks operating in the Middle East and North Africa (MENA). They concluded that banks in the MENA region exhibited reasonable levels of efficiency and that ownership structures (foreign vs. domestic) had no significant effect on efficiency. Their study, however, uses the intermediation approach and makes no cross-country comparison of efficiency.

Of particular interest to our research are comparative efficiency studies between Islamic banks and conventional banks. Al Jarrah and Molyneux (2003) investigate the efficiency of 82 banks in selected MENA countries – Jordan, Egypt, Saudi Arabia and Bahrain, over the period 1992-2000 using a stochastic frontier approach. Islamic banks are found to be the most cost and profit efficient and investment banks the least efficient, Bahraini banks are found to be the most cost and profit efficient.

Al-Shammari (2003), investigates the cost efficiency of 72 banks operating in the GCC countries over the period 1995-1999 using a stochastic frontier approach. He finds Islamic banks

to be the most cost efficient and investment banks the least efficient, with 91% and 84% average efficiency scores, respectively. Overall, Saudi banks are found to be the most cost efficient, followed by UAE banks, with Qatari banks being least efficient (92%, 90%, and 83% respectively).

Studies outside of the GCC/MENA region lead to similar conclusions. A study by Abdul Majid et al. (2003) of Malaysian banks over the period 1993-2000 reveal Islamic banks to be more efficient, albeit marginally, than conventional banks. In a more recent study, Sufian (2007) finds domestic Islamic Malaysian banks to be more efficient than foreign Islamic banks operating in Malaysia. No comparison, however, was made with their conventional counterpart. El-Gamal and Inanoglu (2005) find Islamic banks to be among the most efficient banks operating in Turkey.

The contribution of the current study to the literature is that: 1. it, for the first time, isolates the banking system in the GCC countries in one study, and 2. it investigates the pattern of efficiency change on individual country, as well as on cross-country bases. The latter helps the banks, and the policy makers, to be aware of the effects of their decision on the prospect of future survival in an ever-growing competitive environment.

METHODOLOGY

Examination of the hypotheses set out above requires relative efficiency of banks performance over time. The (relative) efficiency is generally represented by production functions, which can be estimated and evaluated by one of the two widely used parametric and nonparametric approaches of Stochastic Frontier Analysis (SFA) and Data Envelopment Analysis (DEA), respectively. The two approaches use different techniques to envelop data more or less tightly. To that end, they make different accommodations for random noise and for flexibility in the structure of production technology. It is these two different accommodations that generate the strengths and weaknesses of the two approaches.

While DEA involves mathematical programming, SFA is based on econometric methods. This difference in methodology brings about advantages and disadvantages for each of these approaches. Perhaps the most important advantage of the parametric approach is that it has the ability of differentiating between stochastic noise and inefficiency of the production process. This advantage is, however, offset by the SFA's requirement for having a pre-determined production function, common between all the DMUs (decision making units). In addition, the SFA approach can handle only one variable as input. In cases where there are n inputs (n>1), normalization of n-1 inputs is required to make SFA analysis possible. Further, the SFA technique does not allow for direct breakdown of overall efficiency in its various components of technical, scale, and pure efficiency, and comparing each components' changes over time. This breakdown is made possible through Malmquist productivity index (MPI), developed by Malmquist (1953), which is based on DEA non-parametric technique. As far as this study is concerned, a pre-determined common production function for all 6 GCC banking industries seems problematic, which would jeopardize the validity of results. Further the three inputs used in this study require a technique that can handle

multiple input problems. And finally, the use of Malmquist productivity index approach along with the DEA make the analysis harmonized across the study. For this reason we use the DEA in this study. The following is a brief outline of the DEA and Malmquist techniques.

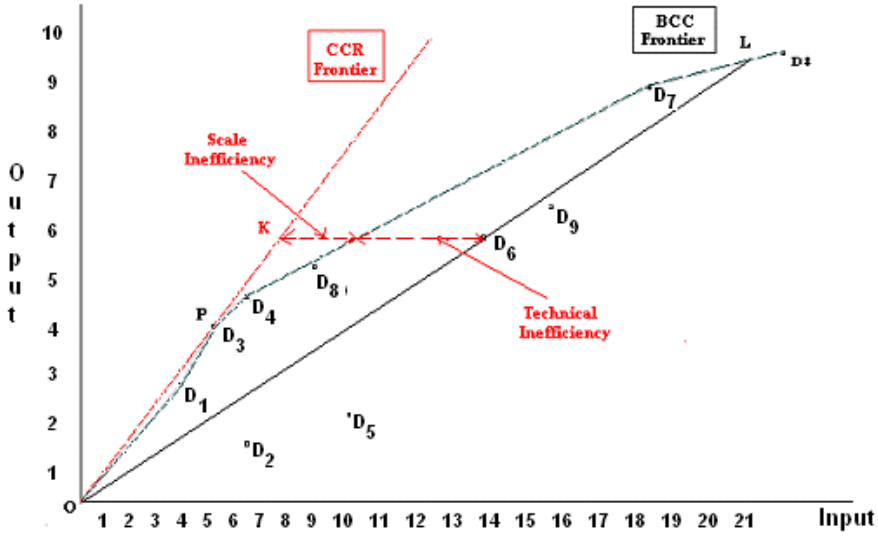
Data Envelopment Analysis

The mathematical approach to construct production frontiers and measure efficiency relative to these constructed frontiers is called data envelopment analysis (DEA). The non-parametric approach of DEA calculates a discrete piecewise frontier determined by the set of Pareto-efficient decision-making units, DMUs - in our case banks. This results in an understanding about each DMU's stand in relation to the rest of the DMUs. This is in contrast to the focus on the averages which is the case with parametric, statistical approaches.

The principle of DEA lies in the definition of efficiency and productivity: if a given bank, C, is capable of using X(C) input to produce Y(C) unit of output, then other banks should also be able to do the same if they were to operate efficiently. Similarly, if bank B is capable of using X(B) units of input for Y(B) unit of outputs, then other producers should also be capable of the same production schedule. Banks B, C and others can be combined to form a composite (virtual) bank with composite (virtual) inputs and composite (virtual) outputs. The non-efficient banks will be compared with this set.

As mentioned earlier, DEA does not require any assumption about the functional form of the production function in question. The only requirement in this approach is that each DMU lies on or below the frontier. Each DMU not on the frontier is scaled against a convex combination of the DMUs on the frontier facet closest to it. For each inefficient DMU, DEA identifies the sources and level of inefficiency for each of the inputs/outputs, depending on the orientation of the study. The level of inefficiency is determined by comparison to a single reference DMU or a convex combination of other referent DMUs located on the efficient frontier.

DEA consists of a variety of models, each one suitable for different settings. The two most commonly applied models are the BCC and CCR models. The CCR model (developed by Charnes, Cooper, and Rhodes, 1978) identifies the source, and estimates the amount, of inefficiencies, as well as yielding an objective evaluation of overall efficiency. The BCC model (developed by Banker, Charnes, and Cooper) distinguishes between technical and scale efficiencies by estimating pure technical efficiency at the given scale of operation, and identifying whether increasing, decreasing, or constant returns to scale possibilities are present for further exploitation. While the BCC model identifies the most efficient banks, the CCR model identifies the most productive banks. This is why more DMUs are introduced as efficient by the BCC model than the number of productive DMUs identified by the CCR model. The following figure gives a geometrical presentation of the two models.



DEA could be applied after deciding on the orientation of the study, which identifies whether the decision-makers have more say in determining the level and/or the type of inputs, or the level of output. The results of an input-oriented DEA indicate the degree of inefficiency in the application of inputs, holding the output fixed. The output-oriented DEA indicates the relative efficiencies of the production process of each decision making unit. The orientation of the study depends on the aim of the study and/or the nature of the industry in question. In the current study, we are seeking the cost efficiency of the banks, which moves us towards an input-oriented analysis.

The input-oriented BCC model is structured as:

Minimize: $Z_0 = \theta - \epsilon.\beta.S^+ - \epsilon.\beta.S^-$ (1)

Subject to:

$$Y\lambda - S^+ = Y_0$$
$$\theta X_0 - X\lambda - S^- = 0$$
$$\beta.\ \lambda \geq 1 \quad (2)$$
$$\lambda,\ S^+,\ S^- \geq 0$$

Here λ is a N x 1vector of constants, Y and X are the output and input matrices, respectively, S is the slack variables, β is a (1× N) row of 1s, and θ is a scalar the value of which will be the efficiency score of the i^{th} DMU. The constraint of ß. $\lambda \geq 1$ allows for variable return to scale of operation. The presence of the non-Archimedean ε allows the minimization over θ to preempt the optimization involving the slacks.

Efficiency over Time

We can include the time component into the production function and measure and analyze the effects of time on the improvement of performance. This is because the production technology of almost all industries is subject to modification based on state-of-the-art technology in the industry in question. Hence the time component is in fact presenting the change in response to production technologies. With this, the production function takes the more challenging form of $Y = f\ (X_1, X_2, \ldots\ldots, X_n, Z_1, Z_2, \ldots.., Z_n, e, t\ )$, where t denotes the element of time, or in fact technology. To take account of the components of the efficiency changes over time we use the Malmquist Total factor productivity index (TFP) introduced by Malmquist (1953). According to Sufian (2006) three advantages differentiates Malmquist productivity indices (MPI) over the alternative approaches (Tornqvist and Fisher indices): Firstly, it does not require specific assumption about the profit and cost function. Secondly, it does not require data on input/output prices. Thirdly, it enables the decomposition of productivity change into technical efficiency change and changes in the best practice frontier (technical/technological change) even if only panel data is available. The shortcoming of the technique is, however, that it needs to compute the distance function. This disadvantage is solved by using DEA.

The input oriented MPI of the i^{th} DMU is formulated as:

$$MPI_i^{t+1}\left(y^{t+1}, x^{t+1}, y^t, x^t\right) = \left[\frac{D_i^t(y^{t+1}, x^{t+1})}{D_i^t(y^t, x^t)} \times \frac{D_i^{t+1}(y^{t+1}, x^{t+1})}{D_i^{t+1}(y^t, x^t)}\right]^{1/2} \quad (3)$$

Where is a distance function measuring the efficiency of the process of producing y^t output using x^t input in the period t. Equation (3) could also be reformulated as:

$$MPI_i^{t+1}\left(y^{t+1}, x^{t+1}, y^t, x^t\right) = \underbrace{\frac{D_i^{t+1}(y^{t+1}, x^{t+1})}{D_i^t(y^t, x^t)}}_{\text{EfficiencyChange}} \times \underbrace{\left[\frac{D_i^t(y^{t+1}, x^{t+1})}{D_i^{t+1}(y^{t+1}, x^{t+1})} \times \frac{D_i^t(y^{t+1}, x^{t+1})}{D_i^{t+1}(y^t, x^t)}\right]^{1/2}}_{\text{TechnolgyChange}} \quad (4)$$

If the MPI (or any of the two components) is greater (less) than 1, it indicates progress (regress).

The first expression in (4) measures the change in i^{th} DMU's position relative to the frontier over the two periods: other things equal, it measures how the i^{th} DMU's relative efficiency has changed between the two periods. In Malmquist literature this is called efficiency change, or the "catching up effect". The second expression contains a geometric average of two alternative measures of shifts in the technology frontier between period "*t"* and period "*t+1"*. Firstly, the shift is measured in relation to year "*t"* observations. Secondly, the shift is measured in relation to year "*t+1"* observations. This shift measures the technological change over the two period.

Technological progress is defined as the progress of the ideas that specifies all activities of an organization towards gaining economic value, or profit. It is changes in the knowledge about product technologies, as well as knowledge about the process by which goods and services are produced. It also comprises the progress in the specification of how activities are organized. In macroeconomic level, this is the most important factor insuring the economic growth of a country. At a micro level, it has the same effect on the progress of an organization (Lipsey and Carlaw, 2004).

The original MPI is based on the assumption of constant returns to scale. This assumption causes overestimation (underestimation) of productivity if the production process demonstrates decreasing (increasing) returns to scale. To solve this problem Fare *et al* (1994) suggest generalized MPI that includes scale index to take account of the effect of economies of scale on productivity.

Another efficiency change, namely the index of variable returns to scale (VRS) efficiency change, is the ratio of the VRS technical efficiency of period $t+1$ to that of period t. This also could be extracted from MPI. Moreover, the index of scale efficiency change of a DMU can be calculated as the ratio of its scale efficiency of period $t+1$ to that of period t (For a thorough study of Malmquist Productivity Index and its decomposition see for example Lovell, 2003)

DATA AND MODEL SPECIFICATION

Data

Balance sheet and income and expenses data of GCC banks were extracted from the BankScope database for the period 2000-2005. For missing data, whenever possible, individual bank annual reports were used. Out of 135 GCC banks reported in the BanScope database, only 56 banks (41 conventional banks and 15 Islamic banks) consistently reported the input and output variables needed for the analysis between 2000 and 2005. Development banks, mortgage banks, and special finance houses were excluded from the analysis. The UAE had the largest number of banks in the sample: 14 conventional and 3 Islamic banks (see Tables 2-7) while Bahrain had the largest number of Islamic banks (5). Oman had no full-fledged Islamic banks, but offered Islamic banking services through the Islamic windows of its conventional banks.

Selection of Input and Output

One of the most important issues in analyzing the efficiency of banking sector has proven to be the choice of input and output variables. This choice depends on the view held about the banking production function. Little agreement exists as to what a bank produces or how to measure its output. However, the two approaches of “intermediation” and “production” have been dominant in most of the mentioned studies.

The Intermediation approach views banks as mediators between the supply and the demand of funds, and as such they take deposits and other purchased funds as inputs and transforms them into loans and other assets, as outputs. The Production approach, on the other hand, views banks as producers of loans and deposits using different inputs such as labor and capital. At present, there is no agreement on the explicit definition and measurement of inputs and outputs of the industry. More precisely, researchers have always found significant difficulties in the definition and measurement of the concept of bank output (Altunbas and Molyneux, 1996). However, if we differentiate between input and output in a way that, ceteris paribus, if it is desirable to increase the quantity of the variable, it is an output, and if it is undesirable to have an increase in its value, it is an input (Charnes, Cooper, and Rhodes, 1978), then the intermediation approach become problematic as no bank would wish to have less deposits. It is for this reason that we adopt the production approach.

Accordingly, for the definition of inputs and outputs, we adopt a variation of the production approach used by Ferrier and Lovell (1990) and Cebenoyan and Register (1989) and conform more to a recent study by Ozkan-Gunay and Tektas (2006) about the Turkish banking industry (see Table1). We use three inputs and three outputs in our model. The three inputs represent resources/expenses required to operate a bank. These are: personal expenses measured by salary, other operating expenses (including expenses incurred for premises and fixed assets), and interest expense (or return to depositors in the case of Islamic banks). The three outputs represent desired outcomes. These are: earning assets, total deposits, and operating income. Banks allocate resources and control expenses and strive to maximize earning assets and income. Banks that do this better than their peers fall on the efficient frontier. Banks that utilize too much input for a given level of output, or produces too little output for a given level of input relative to their peers are considered inefficient and will fall below the efficient frontier.

Table 1: Selected Efficiency studies and choice of Input-Output Variables

Author	Input	Output		Approach	Country	Period
Ozkan-Gunay and Tektas (2006)	- Personnel expenses	- Total deposits	DEA	mix	Turkey	1990-2001
	- Administrative expenses	- Total loans				
	- Interest expenses	- Total securities				
		- Interest income				
		- Non-interest income				
Bos et al (2005)	- Price of Fixed Assets	- Interbank Loans	SFA	Inter-mediation	Germany	1993-2003
	- Price of Labor	- Customer Loans				
	- Price of borrowed Funds	- Securities				
	- Equity					

Table 1: Selected Efficiency studies and choice of Input-Output Variables						
Author	Input	Output		Approach	Country	Period
Matousek R. and Taci A. (2005)	- Employee salary	- Loans, investments and securities		Distribution Free Approach	Czech Republic	1993-1998
	- Price of Capital	- Demand deposits				
	- Price of funds					
Weill (2004)	- Personnel expenses	- Loans	SFA, DFA, DEA	Inter-mediation	France, Germany, Italy, Spain, Switzerland	1992-1998
	- Other non-interest expenses	- Investment assets				
	- Interest paid					
Yamori et al (2003)	- Average number of employees	- Loans & bills discounted	SFA	Inter-mediation	Japan	2002
	- Capital	- Securities & trading account securities				
	- Deposits	- Cash, cash due from banks and call loans				
Barr et al. (2002)	- Salary expense	- Earning assets	DEA	Inter-mediation	US	1984-1998
	- Premises & fixed assets	- Interest income				
	- Other non-interest expenses	- Non-interest income				
	- Purchased funds					

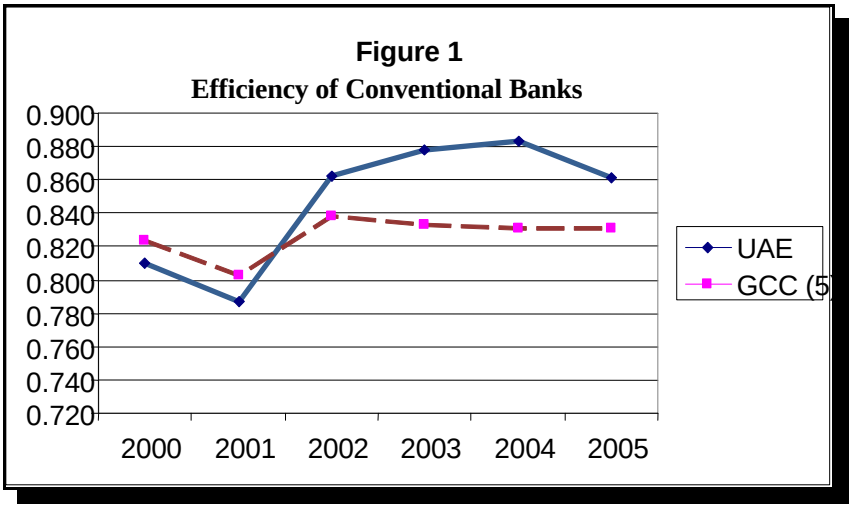
Table 1: Selected Efficiency studies and choice of Input-Output Variables						
Author	Input	Output		Approach	Country	Period
Jemric & Vujcic (2002)	- Labor	- Loans	DEA	Inter-mediation	Croatia	1995-2000
	- Capital	- Securities				
	- Deposits					
Ferrier & Lovell (1990)	- Total number of employees	- Number of demand deposit accounts		Production	USA	1984
	- Occupancy costs and expenditure on furniture and equipment,	- Number of time deposit accounts				
	- Expenditure on materials	- Number of real estate loans,				
		- Number of installment loans,				
		- Number of commercial loans.				
Source: Authors compilation based on a review of the literature.						

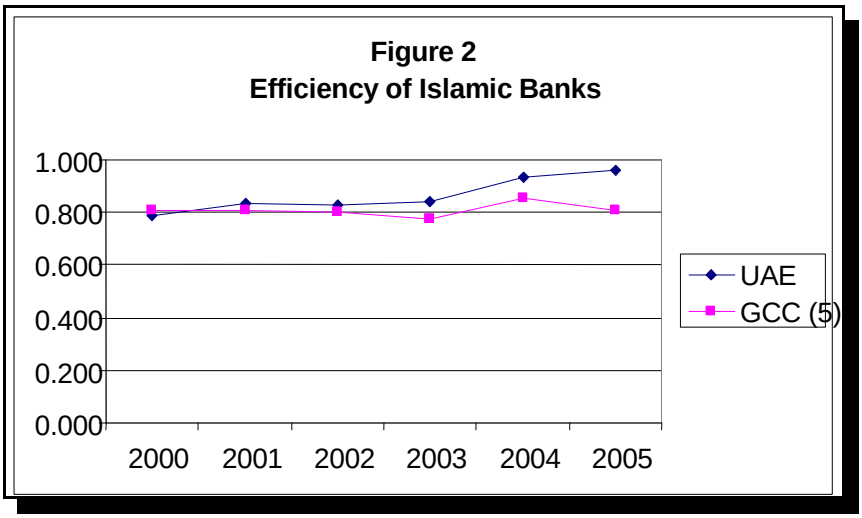
RESULTS

The UAE versus GCC Conventional Banks

The efficiency analyses, presented in tables 2-9 cover the period 2000 – 2005. With 14 conventional, and 3 Islamic banks, the U.A.E. has the largest number of banks in this study, while Bahrain has the most number of Islamic banks, 6, within the GCC.

Figures 1 and 2 show the movements of the efficiencies of the conventional and Islamic banks of the UAE and GCC (UAE excluded). The UAE conventional banks experience a high correlation (just over 87%) with that of the remaining GCC countries, while the degree of the correlation between the Islamic banks in the UAE and in the rest of the GCC countries does not exceed 51%.





In numerical terms, the UAE conventional and Islamic banks had a better performance than the average GCC banks except for 2000 and 2001 for conventional banks and 2000 for Islamic banks. Looking at UAE conventional banks, only the Kuwaiti and Qatari banks have had a superior efficiency than the UAE conventional banks. The numerical superior efficiency of these banks over the UAE conventional banks, and the latter over the remaining GCC banks are not supported by statistical tests which claim that the banks in GCC countries, on a cross-country basis, have no supremacy over each other. The same is true when comparing the GCC Islamic banks on a cross-country basis.

In the UAE, Abu Dhabi Commercial Bank and National Bank of Dubai constantly appear on the best practice frontier. In fact, both banks prove to be not only two of the most efficient

conventional banks, but also the most productive banks as well. In 2005, the National Bank of Abu Dhabi and the Bank of Sharjah joined the two banks as the best practices. Future study would show whether the newcomers could safeguard their new status or not.

Another outstanding point is the decreasing number of UAE conventional and Islamic bank showing increasing returns to scale production possibility. Bearing in mind that institutes with this characteristic are the prime target for take-over, the UAE banks seem to be moving fast away from this characteristic. Leading the way are the Islamic banks that have shown the ability to produce on the favorable scale of constant returns.

While previous studies have consistently shown the higher efficiency of GCC Islamic banks over their conventional counterparts (Hussein, 2004; Al-Jarrah and Molyneux, 2003; Al-Shammari, 2003), the analyses show inconclusive results concerning the numerical efficiencies of Islamic banks versus conventional both at the UAE and the GCC levels, although the UAE Islamic banks did score better performance during the last two years of the study. The statistical tests show no significant differences between the performance of the Islamic banks and their conventional counterparts.

Table 8 indicates the amount of wastage in each input which could have been saved had the banks in question operated at the level operated by the efficient firms in this study. According to the results, the GCC countries, as a whole, have wasted close to $570 million (real the year 2000 money) due to the inefficiency incurred during the 6-year period under study. In this, the share of UAE 17 banks is approximately 20% of the total wastage. The waste per bank is $7.81 million for the UAE and $17 million for the rest of the GCC countries. The majority of the wastes incurred in "other operating expenses" while the "interest paid" takes the lowest waste.

Both in the UAE and the remaining of the GCC countries the Islamic banks appear to have wasted less compared to the conventional banks. In the UAE, Islamic banks are accountable for not more than 14% of the wastage incurred during this period, while the figure for the Islamic banks of the other GCC countries is approximately 25%.

Efficiency over Time

As discussed earlier, it is possible to breakdown the total factor productivity (TFP) changes over the years under study into technical efficiency and technological efficiency change, and obtain indices for pure efficiency and scale efficiency change over the years under study. Here we briefly discuss the changes in the mentioned factors. Table 9 shows the average annual changes in the TFP of the GCC banks.

Technical efficiency (catching up effect)

The technical efficiency, or the catching-up effect, is the decision making unit’s ability to reverse the inefficiency observed in the previous period and “catch-up” with the best practices of the former period.

The Islamic UAE banks gained efficiency throughout this period (except for 2002 when the efficiency gain was – 0.5%). This left the UAE Islamic banks with an average gain of 4.8% over the period under study. In the other GCC countries, on the other hand, the losses incurred in 2002, 2004, and 2005 compared to the prior year left these countries with an average of 0.4%.

Country-wise, the UAE Islamic banks experienced the best average efficiency gain amongst the entire GCC Islamic banks by an annual average of 4.8%. They were followed by Kuwait (with 3.7% average annual efficiency gain) and Qatar (with 2.5% average annual efficiency gain). With a mere 0.7% average annual gain, the UAE conventional banks stand the 4th amongst the GCC countries, where Qatar leads the average annual gain (2.1%) followed by Kuwait (1.7%).

The UAE conventional banks experienced their highest efficiency loss of close to 5% in 2001 and 2005 relative to the 2000 and 2004 frontier, respectively. Their highest efficiency gain incurred in 2002, when the UAE conventional bank gained over 11% of average efficiency relative to their score of 2001. In average, the UAE experienced a mere 0.7% over the years under study. The remaining GCC conventional banks did, in average, more or less the same as the UAE conventional banks with an average efficiency gain of 0.5% over the period under study. For these banks too the years 2002 and 2005 brought about some efficiency gain which was almost cancelled out but the losses in other years.

Technical Efficiency Change (Technological Change)

The technical change is the managements' ability to make the best use of the technological advances and/or allocating resources with the least possible waste.

The UAE banks have an average annual technical efficiency gain of 6.8 %. In this, the conventional banks with 7.3% lead the Islamic bank with annual average technical efficiency gain of 4.6 %. As a matter of fact, the conventional banks of the UAE managed to have technical efficiency gain throughout the period under study. The highest technical efficiency gain for both banks took place in 2002 when they experienced 13.6% technical efficiency gain over that of the previous year. The Islamic banks of the UAE also gained technical efficiency until 2004. The Islamic banks too saw the highest increase in 2002, when the technical efficiency gain was just above 28%. However, with 11% and 17% technical efficiency loss in 2004 and 2005, respectively compared to that of the prior years, the average annual technical efficiency gain of the UAE Islamic bank was reduced to a mere 4.6% gain. The figure for conventional banks of the country is over 7% average annual gain.

Between 2001 and 2003, the UAE Islamic banks experienced the highest technical growth amongst the GCC Islamic banks. However, when the paths of improvement were inversed in 2004 and 2005, it was Kuwaiti and Bahraini banks, respectively, who took over. However, with an average annual technical improvement of 9.3%, the Qatari Islamic banks lead the other Islamic banks in the region.

The UAE conventional banks, on the other hand, lead the other GCC conventional bank with an annual average technical efficiency of 7.3%. The UAE conventional banks are followed by the Omani and Saudi conventional banks with annual averages of 6.1% and 6.0%, respectively. Overall (Islamic and conventional banks combined), it is the UAE banks that had the highest technical efficiency gain amongst the GCC countries.

Total factor productivity change

The results show that over the period of study, the TFP of both the GCC banks (excluding the U.A.E) and the UAE has experienced some annual average gains. The GCC (5) Conventional banks have gained, in average, 5.79% of TFP while the Islamic banks have had an average increase of 6.4 %. The average TFP gain for the UAE conventional and Islamic banks were 8.1% and 9.5%, respectively. This shows a better improvement in UAE banks' performance compared to the average improvement in the remaining of the GCC countries.

As far as the UAE is concerned, the managerial skill in efficient use of technical and technological advancement, measured by the technical change has the highest share in the TFP's improvement.

The other GCC Islamic banks, however, suffered the most from what proved to be the UAE Islamic banks strength, namely the managerial skill to catch-up with the past inefficiencies. These banks lost around 0.4% (average annual) in efficiency while their UAE counterparts gained over 4% annual average in that. We will soon have a closer look at this phenomenon.

Pure Efficiency Change

Typically, pure efficiency gain reflects the managements' efforts in bettering investments in organizational factors, such as a better balance between input and output.

Although both Islamic and conventional banks gained less pure efficiency than the other two discussed above, the end result was average annual pure efficiency gains of 0.9% and 4.7% for conventional and Islamic banks, respectively.

The year 2004 seems to be the best year for UAE Islamic banks with regard to pure efficiency gain (gaining 11.4% relative to 2003), while the best year for the UAE conventional banks was 2002, marking the highest pure efficiency gain of 10.4% compared to that of 2001.

While the Islamic banks gained pure efficiency every year of the period in question, the conventional banks of the country actually lost efficiency in 2005 and 2001.

With a mere 2.3% and 1.2% annual average, respectively, Qatar and Kuwait lead the GCC conventional banks in gaining pure efficiency. However, it is the UAE Islamic banks who had the highest annual average pure efficiency gain amongst all GCC Islamic banks. Thanks to the Islamic banks' performance, the UAE could follow Qatar in gaining the highest overall annual average in pure efficiency.

Scale Efficiency

Scale efficiency refers to the potential productivity gain from achieving optimal size of firm. In other words, it determines whether the firms can be more productive by increasing the size of their operation.

Surprisingly, the highest scale efficiency gain of the UAE Islamic and conventional banks were 2% and 1%, respectively, both occurring in 2004. During the other years under study, the efficiency change (gain or loss) was insignificant (less than a fraction of a percent) for both types of the banks. This comes especially surprising considering the highest ever oil price during this period which would, supposedly cause inflow of oil-dollars to the UAE banks.

Even more surprising is the fact that none of the other GCC country banks (either Islamic or conventional) experienced more than 3.5% scale efficiency gain throughout this period. As a matter of fact, the highest scale efficiency gain incurred in Kuwait in 2003, when the Islamic banks could gain, in average, 3.5% scale efficiency gain. Even with this, the performance of the Kuwait banks in other years left the country with a mere 0.4% overall annual average scale efficiency gain.

SUMMARY AND CONCLUDING REMARKS

We believe this paper has enriched the literature in at least three dimensions. First, to our best knowledge, this is the first study that has investigated production efficiency of GCC banks. We have adopted the non-parametric approach of variable returns to scale data envelopment analysis to perform the task. Second, by estimating the Malmquist productivity index for these banks, we were able to decompose banks' total factor productivity into efficiency and pure efficiency, technical, and scale efficiency and monitor gains in these two components over the 2000-2005 time frame, also a first for this area. Finally, this study is also the first one to test for the difference in production efficiency between Islamic and conventional banks.

The study has adopted an input orientation approach of DEA, with "personnel expenses", "other operating expenses", and "interest expenses /return to depositors" as three inputs. Our results indicate that all the input variables play significant role in determining the relative efficiency of GCC banks. Results further indicate that a relatively large number of GCC banks appear on the best practice frontier, which could partly be due to the small number of banks relative to the sum of number of inputs and outputs.

When looking at the overall efficiency, UAE banks (Islamic and conventional) reported higher efficiency scores then their GCC counterpart most of the time, but this difference in efficiency was not found to be statistically significant. Only the Kuwaiti and Qatari banks had a superior efficiency over UAE conventional banks, still not statistically significant. The same held true when comparing GCC Islamic banks on a cross-country basis.

While previous studies have consistently shown the higher efficiency of GCC Islamic banks over their conventional counterparts, the analyses show inconclusive results concerning the

comparative numerical efficiencies of Islamic versus conventional banks both at the UAE and the GCC levels, although the UAE Islamic banks consistently recorded higher efficiency scores (except for the year 2000) over time when compared to Islamic banks in the other GCC countries.

The Malmquist productivity indices show that banks in the GCC, whether looking at Islamic banks, conventional banks or all banks, recorded a meager 0.6% total productivity gain over the years of study. Overall (Islamic and conventional banks combined), it is the UAE banks that had the highest technical efficiency gain amongst the GCC countries. UAE Islamic banks also had the highest annual average pure efficiency gain amongst all GCC Islamic banks. This gain was recorded at a time when all banks recorded either a loss or an insignificant gain (less than 1%) in pure efficiency. In the area of scale efficiency the gain recorded by all banks, conventional and/or Islamic, was very small. The maximum gain recorded was by Conventional banks outside of the UAE (0.5%). UAE Islamic banks had 0.1% gain in scale efficiency, while UAE conventional ones had a loss of 0.1%.

BIBLIOGRAPHY

Abdul Majid M., Nor, N. & Said, F.F. (2003). Efficiency of Islamic Banks in Malaysia. *Proceedings of the Fifth International Conference on Islamic Economics and Finance*. Bahrain

Al-Jarrah, I. & P. Molyneux, (2003). Efficiency in Arabian Banking. *Proceedings of the International Conference on Financial Development in Arab Countries.* Abu Dhabi, UAE.

Al-Shammari, S. (2003*). Structure – Conduct – Performance & Efficiency in Gulf Cooperation Council (CCC) Banking Markets*. Unpublished Doctoral Dissertation, Bangor University

Altunbas, Y. & P. Molyneux (1996). Economies of Scale and Scope in European Banking. *Applied Financial Economics*, 6(4), 367 – 375.

Athanassopoulos A. & D. Giokas (2000). The use of Data Envelopment Analysis in Banking Institutions: Evidence from the Commercial Bank of Greece. *Interfaces* 30(2) ,81-95.

Ayadi F., A. Adebayo & E. Omolehinwa (1998). Bank Performance Measurement in a Developing Economy: An Application of Data Envelopment Analysis. *Managerial Finance*. 24(7), 5-16.

Banker R., H. Chang H. & S. Majumdar (1996). Profitability, Productivity, and price Recovery Patterns in the U.S. Telecommunications Industry. *Review of Industrial Organization,* 11(1), 1-17.

Banker R., Chang H. and Natarajan R. (1999). Statistical Tests of Allocative Efficiency Using DEA: An Application to the U.S. Public Accounting Industry. *Proceedings of the Sixth European Workshop on Efficiency and Productivity*, KVL, Copenhagen, Denmark.

Banker R., A. Charnes A. & W.W. Cooper (1984). Some Models for Estimating Technical and Scale inefficiencies in Data Envelopment Analysis. *Management Science,* 30(9), 1078-1092.

Barr R.S., K.A. Killgo, T.F. Siems & S. Zimmel (2002) Evaluating the productive efficiency and performance of U.S. commercial banks. *Managerial Finance* .28(8), 3-25.

Barr R., L. Seiford & T. Siems (1994). Forecasting Bank Failure: A Non-Parametric Approach. *Recherches Economiques de Louvain,* 60(4), 417-429.

Barros C. & M. Garcia (2006). Performance Evaluation of Pension Funds Management Companies with Data Envelopment Analysis. *Risk Management and amp Insurance Review*, 9(2), 165.

Berg, S.A., F. Forsund & E. Jansen E (1991). Technical Efficiency of Norwegian Banks: A Nonparametric Approach to Efficiency Measurement. *Journal of Productivity Analysis*, 2(2), 127-142.

Berg S.A., F. Forsund, L. Hjalmarsson & M. Suominen (1993). Banking Efficiency in the Nordic Countries. *Journal of Banking and Finance*, 17(2), 371-388.

Bos J.W.B, F. Heid, M. Koetter, J.W. Kolari & C.J.M. Kool (2005). Inefficient or just Different? Effects of. Heterogeneity on Bank Efficiency Scores. *Deutsche Bundesbank Discussion Paper Series2:* Banking and Financial Studies No. 15.

Brown K. & M. Skully (2006). Evaluating cost performance of banks in the Asia Pacific *Economic Papers, Special edition*, December.

Camanho A. S. & R. G. Dyson (2005). Cost efficiency, production and value-added models in the analysis of bank branch performance. *Journal of the Operational Research Society*, 56, 483.

Camanho A. S. & R. G. Dyson (1999). Efficiency, size, benchmarks and targets for bank branches: an application of data envelopment analysis. *Journal of the Operational Research Society*, 50, 903-915.

Cebenoyan A.S., & C.A. Register (1989). Banking Efficiency: Stochastic Cost Frontier Analysis Approach, *Proceedings of the 19th Annual FMA Meeting*, Boston, USA.

Charnes A., W. Cooper W. & E.L. Rhodes (1978). Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research*, 2(6), 429-444.

Chen T. & T. Yeh (2000). A Measurement of Bank Efficiency, Ownership and Productivity Changes in Taiwan. *The Service Industries Journal*. 20 (1).

El-Gamal M.A. and H. Inanoglu (2005). Inefficiency and Heterogeneity in Turkish Banking: 1990-2000. *Journal of Applied Econometrics*, 20(5), 641.

Fare R., B. Grosskopf, M. Norris M., & Z. Zhang (1994). Productivity growth, technical progress, and efficiency change in industrialized countries. *American Economic Review*, 84 (1), 66-83.

Fecher F. & P. Pestieau (1993). Efficiency and Competition in O.E.C.D. Financial Services in H.O. Fried, C.A.K. Lovell, & S. Schmidt, (Eds), The Measurement of Productive Efficiency: Techniques and Applications (pp. 374-385). Oxford University Press.

Ferrier G, & C.A.K Lovell (1990). Measuring Cost Efficiency in Banking: Econometric and Linear Programming Evidence. *Journal of Econometrics*, 46(1), 229-245.

Fukuyama H. (1995). Measuring Efficiency and Productivity Growth in Japanese Banking: A Nonparametric Frontier Approach. *Applied Financial Economics*, 5(2), 95-107.

Grigorian D. & V. Manole (2006). Determinants of Commercial Bank Performance in Transition: An Application of Data Envelopment Analysis. *Comparative Economic Studies*, 48(3), 497-522.

Hussein, K.A. (2004). Banking Efficiency in Bahrain: Islamic vs. Conventional Banks. *Islamic Research and Training Institute*. Islamic Development Bank. Jeddah.

Jemric I. & B. Vujcic (2002). Efficiency of banks in Croatia: A DEA approach. *Comparative Economic Studies*. 44(2/3), 169-193.

Kirkwood J., & D. Nahm (2006). Australian Banking Efficiency and its Relation to Stock Returns. *The Economic Record*, 82 (258), 253-267.

Krishnasamy G, Hanuum Ridzwa A., and Perumal V. (2003). Malaysian Post Merger Banks' Productivity: Application of Malmquist Productivity Index. *Managerial Finance*, 30(4), 63-74.

Levine, R. (1998). The legal Environment, Banks, and Long-Run Economic Growth. *Journal of Money, Credit, and Banking*, 30(3), 596-613.

Levine, R. & S. Zervos, S. (1998). Stock Markets, Banks, and Economic Growth. *The American Economic Review*, 88(3), 537-558.

Lipsey R. & K. Carlaw (2004). Total Factor Productivity and the Measurement of Technological Change. *Canadian Journal of Economics*, 37(4), 1118-50..

Lo S. and W. Lu (2006). Does size matter? Finding The Profitability and Marketability Benchmark of Financial Holding Companies. *Asia-Pacific Journal of Operational Research*, 23(2), 229-246.

Lozano-Vivas A. & J. Pastor. (2006). Relating macro-economic efficiency to financial efficiency: A comparison of fifteen OECD countries over an eighteen year period. *Journal of Prod Anal,* 25(1/2), 5-6.

Malmquist S. (1953). Index Numbers and Indifference Curves. *Trabajos de Estadistica*, 4, 209-42.

Matoucek R. and Taci A. (2005). Efficiency in Banking: Empirical Evidence from the Czech Republic. *Economic Change and Restructuring*, 37(3), 225-244.

Ozkan-Gunay E.N, & A. Tektas, A. (2006). Efficiency Analysis of the Turkish Banking Sector in Precrisis and Crisis Period: A DEA Approach. *Contemporary Economics Policy*, 24 (3), 418-31.

Rao A. (2005). Cost Frontier Efficiency and Risk-Return Analysis in an Emerging Market. *International Review of Financial Analysis*, 14(3), 283-303.

Ruthenberg D. & R. Elias (1996). Cost Economies and Interest Rate Margins in a Unified European Banking Market. *Journal of Economics and Business*, 48(3), 231-49.

Saif I. & H. Yaseen (2005). The Efficiency of Foreign and Domestic Banks in MENA Region: Evidence on Economies of Scale and Scope. *Proceedings of the 12th Annual Conference, Economic Research Forum,* Cairo, Egypt.

Samjeev G.M. (2006). Does Banks' Size Matter in India? *Journal of Services Research*, 6(2).

Sathye M. (2002). Measuring Productivity Changes in Australian Banking: An Application of Malmquist Indices. *Managerial Finance*, 28(9), 48-59.

Sufian F. & M. Abdul Majid (2007). X-Efficiency and share prices in the Singaporean banking sector: A DEA Window Analysis approach. *Investment Management & Financial Innovations*, 4(1), 73-91.

Sufian F. (2007). The efficiency of Islamic Banking Industry in Malaysia: Foreign vs Domestic Banks. *Humanomics,* 23(3), 174-192.

Sufian F. (2006). Size and Total Factor Productivity Growth of Non-Bank Financial Institutions. South Asian Journal of Management, 13(4): 7-24.

Weill L. (2004). Measuring Cost Efficiency in European Banking: A Comparison of Frontier Techniques. *Journal of Productivity Analysis*, 21(2), 133-152.

Yamori N., K. Harimaya & K. Kondo (2003). Are Banks Affiliated with Bank Holding Companies More Efficient Than Independent Banks? The Recent Experience Regarding Japanese Regional BHCs. *Asia-Pacific Financial Markets*. 10(4), 359-76.

Zaire O. (1995). The Effect of Financial Liberalization on the Efficiency of Turkish Commercial Banks. *Applied Financial Economics*, 5(4), 257-64.

Zenios C., Zenios S., Agathocleous K., and Soteriou A. (1999). Benchmarks of the Efficiency of Bank Branches. *Interfaces*, 29(3), 67-83.

Table 2: Efficiency information of 2000							
Year 2000	Average VRS	Significant difference?	Range	Banks on Frontier	Average Scale Efficiency	No. of Banks With IRS	No. of Banks With DRS
Conventional: Overall GCC (41)	0.819 (0.117)	No	0.431 - 1.000	7 (17.1%)	0.962	27 (65.9%)	6 (14.6%)
Islamic: Overall GCC (15)	0.802 (0.182)	No	0.406 - 1.000	5 (33.3%)	0.962	9 (60.0%)	1 (6.7%)
Conventional: GCC Excl. UAE (27)	0.824 (0.126)	No	0.431 - 1.000	5 (18.5%)	0.95	18 (66.7%)	5 (18.5%)
Conventional: U.A.E. (14)	0.810 (0.103)	No	0.601 - 1.000	2 (14.3%)	0.984	9 (64.3)	1 (7.1%)
Islamic: GCC Excl. UAE (12)	0.806 (0.184)	No	0.406 - 1.000	4 (33.3%)	0.956	7 (58.3%)	1 (8.3%)
Islamic: U.A.E. (3)	0.785 (0.212)	No	0.612 - 1.000	1 (33.3%)	0.988	2 (66.7%)	0

Table 3: Efficiency information of 2001							
Year 2001	Average VRS	Significant difference?	Range	Banks on Frontier	Average Scale Efficiency	No. of Banks With IRS	No. of Banks With DRS
Conventional: Overall GCC (41)	0.798 (0.160)	No	0.237 - 1.000	8 (19.5%)	0.967	23 (56.1%)	12 (29.3%)
Islamic: Overall GCC (15)	0.816 (0.201)	No	0.355 - 1.000	5 (33.3%)	0.968	9 (60.0%)	2 (13.3%)
Conventional: GCC Excl. UAE (27)	0.803 (0.169)	No	0.237 - 1.000	5 (18.5%)	0.967	15 (55.5%)	9 (33.3%)
Conventional: U.A.E. (14)	0.787 (0.145)	No	0.567 - 1.000	3 (21.4%)	0.968	8 (57.1)	3 (21.4%)
Islamic: GCC Excl. UAE (12)	0.811 (0.209)	No	0.355 - 1.000	4 (33.3%)	0.963	7 (58.3%)	2 (11.8%)
Islamic: U.A.E. (3)	0.836 (0.201)	No	0.693 - 1.000	1 (33.3%)	0.968	2 (66.7%)	0

Table 4: Efficiency information of 2002							
Year 2002	Average VRS	Significant difference?	Range	Banks on Frontier	Average Scale Efficiency	No. of Banks With IRS	No. of Banks With DRS
Conventional: Overall GCC (41)	0.846 (0.146)	No	0.384 - 1.000	10 (24.4%)	0.969	23 (56.1%)	9 (22.0%)
Islamic: Overall GCC (15)	0.806 (0.220)	No	0.602 - 1.000	6 (40.0%)	0.959	7 (46.7%)	3 (20.0%)
Conventional: GCC Excl. UAE (27)	0.838 (0.164)	No	0.384 - 1.000	7 (25.9%)	0.967	13 (48.1%)	8 (29.6%)
Conventional: U.A.E. (14)	0.863 (0.110)	No	0.602 - 1.000	3 (21.4%)	0.972	10 (71.4)	2 (7.1%)
Islamic: GCC Excl. UAE (12)	0.800 (0.238)	No	0.408 - 1.000	5 (41.7%)	0.953	6 (50.0%)	3 (25.0%)
Islamic: U.A.E. (3)	0.830 (0.161)	No	0.681 - 1.000	1 (33.3%)	0.98	1 (33.3)	0

Table 5: Efficiency information of 2003							
Year 2003	Average VRS	Significant difference?	Range	Banks on Frontier	Average Scale Efficiency	No. of Banks With IRS	No. of Banks With DRS
Conventional: Overall GCC (41)	0.849 (0.152)	No	0.466 - 1.000	11 (26.8%)	0.966	27 (65.9%)	6 (14.6%)
Islamic: Overall GCC (15)	0.789 (0.222)	No	0.323 - 1.000	5 (33.3%)	0.949	9 (60.0%)	2 (13.3%)
Conventional: GCC Excl. UAE (27)	0.833 (0.167)	No	0.466 - 1.000	8 (29.6%)	0.959	17 (63.0%)	4 (14.8%)
Conventional: U.A.E. (14)	0.879 (0.118)	No	0.620 - 1.000	3 (21.4%)	0.978	10 (71.4%)	2 (14.3%)
Islamic: GCC Excl. UAE (12)	0.775 (0.236)	No	0.323 - 1.000	4 (33.3%)	0.942	7 (58.3%)	2 (16.7%)
Islamic: U.A.E. (3)	0.844 (0.179)	No	0.649 - 1.000	1 (33.3%)	0.981	2 (66.7%)	0

Table 6: Efficiency information of 2004

Year 2004	Average VRS	Significant difference?	Range	Banks on Frontier	Average Scale Efficiency	No. of Banks With IRS	No. of Banks With DRS
Conventional: Overall GCC (41)	0.849 (0.164)	No	0.473 - 1.000	14 (34.1%)	0.973	17 (41.5%)	12 (29.3%)
Islamic: Overall GCC (15)	0.869 (0.198)	No	0.437 - 1.000	7 (46.7%)	0.954	4 (26.7%)	5 (33.3%)
Conventional: GCC Excl. UAE (27)	0.831 (0.181)	No	0.473 - 1.000	9 (33.3%)	0.966	12 (44.4%)	9 (33.3%)
Conventional: U.A.E. (14)	0.883 (0.122)	No	0.643 - 1.000	5 (35.7%)	0.987	5 (35.7%)	3 (21.4%)
Islamic: GCC Excl. UAE (12)	0.854 (0.215)	No	0.437 - 1.000	5 (41.7%)	0.943	4 (33.3%)	5 (41.7%)
Islamic: U.A.E. (3)	0.931 (0.120)	No	0.792 - 1.000	2 (66.7%)	1	0	0

Table 7: Efficiency information of 2005

Year 2005	Average VRS	Significant difference?	Range	Banks on Frontier	Average Scale Efficiency	No. of Banks With IRS	No. of Banks With DRS
Conventional: Overall GCC (41)	0.841 (0.168)	No	0.437 - 1.000	13 (31.7%)	0.974	19 (46.3%)	12 (29.3%)
Islamic: Overall GCC (15)	0.841 (0.214)	No	0.413 - 1.000	5 (33.3%)	0.958	5 (33.3%)	6 (40.0%)
Conventional: GCC Excl. UAE (27)	0.831 (0.168)	No	0.437 - 1.000	7 (25.9%)	0.973	13 (48.1%)	10 (37.0%)
Conventional: U.A.E. (14)	0.861 (0.172)	No	0.481 - 1.000	6 (42.9%)	0.976	6 (42.9%)	2 (14.3%)
Islamic: GCC Excl. UAE (12)	0.811 (0.230)	No	0.413 - 1.000	4 (33.3%)	0.95	5 (41.7%)	5 (41.7%)
Islamic: U.A.E. (3)	0.963 (0.033)	No	0.938 - 1.000	1 (33.3%)	0.991	0	1 (33.3%)

Table 8: The amount of lost saving opportunity of inefficient banks (million US dollar, real 2000)					
			Interest Expense	Personnel Expense	Other Operating Expense
2000	UAE	Conventional	0	0	14.21
		Islamic	0	0	1.1
	GCC (Excl. UAE)	Conventional	58.43	4.02	74.74
		Islamic	0	3.06	9.67
2001	UAE	Conventional	0	1.41	5.96
		Islamic	0	0	1
	GCC (Excl. UAE)	Conventional	27.25	5.98	40.75
		Islamic	0	3.59	9.4
2002	UAE	Conventional	0	2.14	11.21
		Islamic	0	0.03	0
	GCC (Excl. UAE)	Conventional	6.48	11.83	39.55
		Islamic	9.33	0	28.03
2003	UAE	Conventional	0	4.12	6.66
		Islamic	9.33	0	28.03
	GCC (Excl. UAE)	Conventional	0	20.51	15.17
		Islamic	0.22	8.13	7.57
2004	UAE	Conventional	0	10.51	2.52
		Islamic	0	0	0
	GCC (Excl. UAE)	Conventional	6.38	6.49	8.76
		Islamic	0	5.49	1.1
2005	UAE	Conventional	17.72	15.96	1.53
		Islamic	6.63	2.43	0
	GCC (Excl. UAE)	Conventional	0	13.52	6.25
		Islamic	0	28.33	0

Table 9: Breakdown of Average Annual Total Factor Productivity changes					
	Efficiency Change	Technical Change	Pure Efficiency Change	Scale Efficiency Change	TFP Change
Overall Average	1.006	1.06	1.004	1.002	1.066
GCC Conventional Average	1.006	1.059	1.003	1.003	1.065
GCC Islamic Average	1.006	1.064	1.007	0.999	1.07
GCC Excl. UAE Overall Average	1.002	1.057	0.999	1.003	1.059
UAE Overall Average	1.014	1.068	1.016	0.999	1.083
GCC Exc. UAE Conventional Average	1.005	1.052	0.999	1.005	1.057
UAE Conventional Average	1.007	1.073	1.009	0.998	1.081
GCC Excl. UAE Islamic Average	0.996	1.069	0.997	0.999	1.064
UAE Islamic Average	1.048	1.046	1.047	1.001	1.095

IS THE CREDIT CRUNCH A SUPPLY SIDE PHENOMENON? A THEORETICAL APPRAISAL

Sudip Ghosh, Penn State University, Berks Campus

ABSTRACT

This paper provides a theoretical understanding of the credit crunch in general and the current crisis in particular. It is still an open question: Does the reduction in equilibrium loan amount result from supply side factors, or is it driven by weak loan demand?

I use a simple micro theoretical model with an uncertain loan customer and an imperfectly competitive bank (with some market power) to explain the current credit situations. The direction of the interest rate is critical in determining whether the credit decline is a consequence of a precipitous drop in loan demand exacerbated by current weakness in the economy or is it due to simple loan reduction by the lending institutions. Comparative static results shows effects of the collateral (borrower), capital adequacy standards of depository institutions, and economic conditions on the loan amount. The link between economic conditions and the loan amount is shown as positive. The connection between collateral and the loan amount is not clear due to presence of limited liabilities and moral hazard. These moral hazard problems may lead to credit rationing by generating a non-monotonic relationship between quoted interest rates and expected rates of return as in the Stiglitz-Weiss model. However, under restrictive conditions the tie between collateral and loan amount is positive.

INTRODUCTION

The recent focus on credit crunch and its link to subprime-mortgage business has put a dark overhang on the U.S. economy. The history of credit crunch is not new to the U.S. market, dating as far back as 1930s followed by episodes in the mid-1970s, early1990s, in 1998, and as recent as 2007. Anecdotal evidence shows the cause for credit crisis as being different for every single period.

The purpose of this paper is to examine the determinants of credit crunch. A theoretical one period static model of the loan market is developed to explain the credit crunch. The loan market comprises of an imperfectly competitive bank (with certain market power) and a loan customer with homogeneous collateral. The model demonstrates how interest rates and capitalization of banks influence bank loans. The interest rate is set by the bank and borrower simply accepts it.

The recent subprime-mortgage turmoil in the U.S. spread around the world, prompting central bank interventions. The worsening credit situation threatens to put more pressure on the housing market, where prices are flat to declining in much of the country. Lenders respond by

tightening credit, making it more difficult and expensive for businesses, households, and government to borrow. According to the Chairman of the Federal Reserve, Ben Bernanke, "Rising delinquencies and foreclosures are creating personal, economic and social distress for many homeowners and communities - problems that likely will get worse before they get better."

During the Great Depression, over forty percent of depository institutions operating at the beginning of the 1930s failed, and bank lending declined drastically resulting in the credit crunch. According to Bernanke the "credit crunch" resulted from central bank dysfunction which ultimately prolonged the depression and subsequent recovery. In 1974, however, sudden supply side disturbances caused oil prices to go up triggering inflation pressure and rise in the interest rates, sending the market into disarray. Loans available to marginal borrowers were drastically cut leading to the 1974 recession and the subsequent credit crunch.

The credit bottleneck once again emerged in the U.S. economy in the early 1990s with the "junk bond" and the Savings and Loans fiasco. The softening of bank regulations in early 1980s, the dramatic increase in depository institution failures between1988 and 1991 and the Federal Deposit Insurance Corporation Improvement Act (FDICIA) of 1991 regenerated interest in depository institution insolvency risk - a determination to keep these risks under control (Berger and Udell 1992). Often mis-pricing of risks by deposit insurers is, in large part, blamed for the banking crisis. During the past twenty years, bank regulators have increasingly adopted risk-sensitive regulations starting with Basel I. The main objective of Basel I is to protect commercial banks from insolvency using risk-related capital requirements. Thus promotion of market discipline along with recapitalization of banks is consistent with public policy ranging from the provisions of FDICIA to Basel Accord II.

The failure of the Long-Term Capital Management (LTCM) in 1998 and its ultimate demise in early 2000 also added a new dimension; problems associated with hedge funds. In August of 1998 Russia defaulted on its debt and the financial markets came unraveled and LTCM nearly went bankrupt. The Federal Reserve Bank of New York sponsored a bailout of LTCM by its creditor banks and justified its intervention to avoid a financial crisis whose effects could have been more severe. This is another face of the credit crunch.

The efforts to discipline depository institutions through regulations lead to reduction of bank loans popularly known as the credit crunch.[1] The new capital adequacy standard, commonly known as the Risk Based Capital requirements (RBC) often squeezed part of banks equity. Thus, the credit crunch became synonymous to the capital crunch. Here the key issues were whether a binding capital constraint (induced by either higher capital standards or a weakened capital base) resulted in reduced bank lending or was it due to precipitous decline in loan demand exacerbated by the 1990-91 recession?

It is important to understand that in a partial equilibrium framework the usual graphical analysis of supply and demand does not necessarily give an equilibrium solution due to the nature of the backward bending credit supply function, especially at high levels of the interest rate. Under such circumstances a new equilibrium alludes to *credit rationing* while facing a competitive credit

market. According to Baltensperger (1978), a borrower's loan demand is often turned down, even if this borrower is willing to pay all the price and non-price elements of the loan contract. The non-price element, such as collateral, may be inadequate for the lender. It is also important to note that credit is not a perfectly divisible good. Thus lending more to an individual loan customer may expose the risk for the lender; therefore, the equilibrium rate may be a nonlinear function of the loan amount. Thus the market becomes inefficient which is consistent with a nonlinear loan function and non-perfect divisibility of loans as pointed out by De Meza and Webb (1992).

Circumstances under which banks are unable to identify the borrowers lead banks to consider sorting devices constructed in such a way that each type of borrower will choose a specific type of contract. Self-selection by clients will result in product differentiation. This idea is to modify bank strategy and offer different contracts. Based on the borrower choice banks will be able to identify the amount of collateral needed for that specific loan and hence the risk. This idea has been explored by Mussa and Rosen (1978) in the case of a monopoly for a durable good and by Bester (1985) who establishes that equilibrium in the credit market exists without credit rationing.

While academicians and practitioners grappled to understand the reasons behind slow credit growth, skeptics argued that the newly initiated RBC requirement caused a reduction in commercial lending where a significant number of borrowers who otherwise would have been funded were denied credit or were priced out of the market. The term "priced out" means that loan prices were set high so that loan customers elected not to obtain such loans. The banks facing binding capital constraints (during periods marked by large loan losses and falling earnings) modified their balance sheets by either issuing new securities (to raise capital) or routinely switching to assets that needed fewer equity, from the ones that needed more (Peek and Rosengren, 1995; Yago, 1991)[2]. The banks often preferred the second method to reduce portfolio risks while maintaining new capital standards. Since we are still unclear about the real causes of the credit crunch, we will use the broadest definition of credit crunch following Bernanke and Lown (1995).

Due to research by Berger and Udell (1994), Haulbrich and Watchel (1993), and Thakor (1993) Wojnilower (1990) believed that financial disintermediation resulted in credit reduction. Gennote and Pyle (1991) analyzed the impact of capital regulations on bank risk and assumed that banks can only invest in projects with positive Net Present Value since banks have the ability to screen and monitor industrial projects. Also, the banks capital-asset ratio should be tied to such investments which in turn would determine the quality of banks' assets evaluated by bank examiners. Shrieves and Dahl (1995) showed that bank portfolios were more sensitive to external shocks in the early 1990s than in the later 1980s. Hancock and Wilcox (1995) hold a similar view that capital shocks appear to have played at least a partial role in the lending decline. Berger and Udell (1991) show partial or no evidence of a decline in lending related activities due to the RBC requirements.

It has also been argued in the literature that emergence of new competition from other financial intermediaries abated the role of banks as leaders in the lending business. New loan loss reserve policies, heightened scrutiny by regulators, and increased enforcement literally forced banks

to substitute into safer assets. These policy changes affected the capital needed for banks to operate and reduced their ability to manipulate the leverage ratios. Others like Kim and Santomero (1988), Keeley and Furlong (1990), Flannery (1991), Gennotte and Pyle (1991), and Avery and Berger (1991) derived the effect of changed capital requirements on bank portfolio and insolvency risk using both theory and empirics. There is no consensus among economists about the structure of the RBC requirements and whether it is partly responsible for credit bottlenecks and the ensuing credit crunch.

THE THEORETICAL MODEL OF LOAN DEMAND

In this section, a simple partial equilibrium one period static model of loan demand is constructed to check for the issue of credit crunch. For convenience, an inverse loan demand: $r^L = g(L, K)$ is assumed to have the following signs $g_L(L,K) < 0, g_K(L,K) < 0$, where r^L is the exogenous interest rate on loans, g is the inverted demand function of loan L and equity K (or collateral) of the loan customer. In the loan market, the bank determines the maximum amount of the loan for given collateral. The equity K does not vary qualitatively.

Assume a loan customer with observable equity K chooses a loan amount L to maximize expected profits in the presence of limited liability making profits ≥ 0. The expected value of the profit function π^F of the loan customer is given as:

$$E(\pi^F) = \int_{p_0}^{z} [pQ - R_L L] f(p) dp \qquad (2.1)$$

Here, price p, the output price, is random with continuous density function $f(p)$ and support $[p_0, z]$. Consider the additive production function $Q = Q(K+L)$ where K is borrowing firm's collateral and L the loan quantity. It is assumed that all financial capital $(K+L)$ is engaged for purchasing inputs. The collateral K of the borrower is observable, but p is a priori unknown. Production function is assumed to be smooth and concave, and that the second order condition holds.[3] R_L, the interest rate on loans is defined as one plus the loan rate $(1 + r_L)$ and is exogenous to the loan customer. In this setup, the loan customer's zero profit condition is written as

$$\pi^F = [pQ - R_L] = 0. \qquad (2.2)$$

Price p_0 is labeled as the critical price where $\pi^F = 0$. When $p < p_0$, results $\pi^F < 0$ but due to the presence of limited liability it is impossible to generate negative profits. The characteristic of the contract determines the value p_0 and it is required for this condition to hold. It is given by

$$p_0 = \frac{R_L L}{Q(K+L)} \tag{2.3}$$

The expected profit for the loan customer becomes

$$E(\pi^F) = \int_{z}^{p_0} 0 f(\varepsilon) d\varepsilon + \int_{p_0}^{z} \pi^F f(\varepsilon) d\varepsilon. \tag{2.4}$$

The first term on the right-hand side becomes zero due to limited liability. She maximizes expected profit $E(\pi^F)$, with respect to L (loan) and, by applying Leibnitz's Rule gives the first order condition

$$\begin{aligned} \frac{dE(\pi^F)}{dL} = & \int_{p_0}^{z} [pQ(K+L) - R_L] f(p) dp + \frac{dz}{dL}[zQ - R_L L] f(z) \\ & - \frac{dp_0}{dL}(p_0 Q - R_L L) f(p_0) = 0 \end{aligned} \tag{2.5}$$

Using equation (2.2), the last term on the right-hand side of equation (2.5) is eliminated. Since z is constant, dz/dL becomes zero which reduces equation (2.5) into equation (2.6):

$$\int_{p}^{z} (pQ'(K + L) - R_L) f(p) dp = 0 \tag{2.6}$$

Equation (2.6) captures, the expected marginal product of input equals the lending rate times the odds of receiving the payment. The objective is to focus on the loan amount and its relationship to the lending rates.

Assume the following notations: $F(p_0)$ denotes the density function on p_0 and is defined by the following:

$$1 - F(p_0) = \int_{p_0}^{z} f(p) dp \tag{2.7}$$

$$G(p_0) = \int_{p_0}^{Z} p f(p) dp \tag{2.8}$$

$$H(p_0)=\frac{G(p_0)}{1-F(p_0)} \quad \text{where} \quad p_0 \equiv \frac{R_L L}{Q} \tag{2.9}$$

Combining equations (2.7), (2.8), and (2.9), and rearranging terms gives

$$Q'(K+L)H(p_0)=R_L \tag{2.10}$$

the expected marginal product of input equals the lending rate times the odds of receiving the payment making equation (2.10) identical to equation (2.6).

Next, show $dH/dp_0 > 0$. Using $H(p_0)$ from equation (2.9), dH/dp_o is given by

$$\frac{dH}{dp_0}=\frac{G'(1-F)+F'G}{(1-F)^2} \tag{2.11}$$

Applying Leibnitz's Rule in equations (2.8) and (2.9), dH/dp_o results in equation (2.12):

$$\frac{dH}{dp_0}=\frac{-p_0 f(p_0)(1-F)+(-(-f(p_0)))G}{(1-F)^2} \tag{2.12}$$

If the numerator is > 0 then dH/dp_o is > 0. This is true as long as $p \geq p_0$ in the interval $[p_0, z]$ and so does the denominator of equation (2.12). Therefore,

$$\operatorname{sgn}\frac{dH}{dp_0}=\operatorname{sgn}\int_{p_0}^{Z}(p-p_0)dp>0. \tag{2.13}$$

For profit maximization the second order condition holds if and only if $dE^2(\pi^F)/dL^2 < 0$. Differentiating equation (2.5) with respect to L and applying Leibnitz's Rule results in (2.14):

$$\begin{aligned}\frac{d^2E(\pi^F)}{dL^2} &= \int_{p_0}^{z} pQ''(K+L)f(p)\,dp+\frac{dz}{dL}(zQ'(.)-R_L)f(z) \\ &\quad -\frac{dp_0}{dL}(p_0-Q'(.)-R_L)f(p_0)\end{aligned} \tag{2.14}$$

It is not obvious whether $d^2E(\pi^F)/dL^2<0$. Due to the assumption of diminishing returns $Q''(.)<0$. The second term on the right-hand side of equation (2.14) vanishes because dz/dL is zero. The other term dp_0/dL on the right-hand side of equation 2.14 and the terms associated with it is difficult to sign under general conditions.[4] However, determining the correct sign for dp_0/dR_L, dp_0/dL, dp_0/dK, and using $p_0=R_LL/Q(K+L)$, it is clear that the second order conditions holds. The signs for dp_0/dR_L, dp_0/dL, and dp_0/dK are as follows:

$$\frac{dp_0}{dK} = - \frac{R_L L\, Q'_k(.)}{Q^2(.)} < 0, \quad \frac{dp_0}{dR_L} = \frac{L}{Q^2(.)} > 0, \quad \text{and} \quad \frac{dp_0}{dL} = \frac{R_L[1- \varepsilon_{L,\,Q(.)}]}{Q(.)} > 0 \ \text{ or, } < 0$$

The above results rely entirely on the shape of the production function which is assumed be smooth and concave and the loan customer operates under certain restrictions.

The influence of K and R_L, on L is derived by totally differentiating equation (2.10) and combining terms. Thus, for given R_L:

$$\frac{dL}{dK} = - \frac{[Q''(K+L)H(p_0)+Q'(K+L)H'(p_0)\frac{dp_0}{dK}]}{[Q''(K+L)\,H(p_0)\; + \; Q'(K+L)H'(p_0)\frac{dp_0}{dL}]} \tag{2.15}$$

$dL/dK<0$ provided the first order condition holds and that is true only if the second order condition is satisfied, which means that the denominator in (2.15) has to be negative. This means that when borrowing firm's have more collateral, they tend to borrow less for two simple reasons: First, borrowers maximize profit by employing K and reducing their dependence on external funds. Typically, external funds are associated with higher transactions costs and more binding constraints about the loan usage having negative effect on borrower's earnings. Second, funds coming from within the firm are less prone to moral hazard issues as owner-representative are more risk averse (Stiglitz and Weiss 1981). In works of Stiglitz and Weiss (1981), and Holmstrom-Tirole (1993) some of the assets are monetary, or reputation related which shows credit market operations in the presence of moral hazard.

For given K, dL/dR_L is expressed as:

$$\frac{dL}{dR_L} = \frac{[1-H'(p_0)Q'(.)\frac{dp_0}{dR_L}]}{[Q''(K+L)H(p_0)+Q'(K+L)H'(p_0)\frac{dp_0}{dL}]} \tag{2.16}$$

Equation (2.16) is negative provided the first order condition holds and the second order condition is met, assuming that the denominator in equation (2.16) is negative. This is a standard result: The inverse relationship between lending rates and the loan amount. However, equation (2.16) is used to explain the current mortgage crisis.

The recent crisis stems from the disruption caused by rising lending rates and the loan amount not to mention the loan quality. Starting around 2003, the U.S. economy saw an unprecedented growth in home ownership partly due to low mortgage rates. However, as rates start to climb and the initial low fixed rates convert to higher adjustable rates, the U.S. residential market experience widespread delinquencies especially in the sub-prime mortgages. Difficulty in primary mortgage is forcing lenders to use extreme caution while tightening loans. The equation (2.16) explains the main reasons of the current credit squeeze.

A SPECIAL CASE

This case is illustrated by using a specific production function and the uniform distribution $f(p) = 1/(z - p_0)$ where p_0 is zero.[5] The new production function is $Q = (K+L)^{\alpha}$, which is additive in nature where $\alpha \in (0,1)$ so that $p_0 = R_L L/(K+L)^{\alpha}$. The new profit function is shown as

$$E(\pi^F) = \int_{p_0}^{z} [p(K+L)^{\alpha} - R_L L] \frac{1}{z} dp \qquad (3.1)$$

Rearranging terms in equation (.9) and changing sides, the following expression for dL/dR_L is obtained:

$$\frac{dL}{dR_L} = \frac{[2 - \alpha \frac{L}{(K+L)}]}{[z\alpha(\alpha - 1)(K+L)^{\alpha - 2} + \alpha R_L \frac{K}{(K+L)^2}]} \qquad (3.2).$$

The numerator of dL/dR_L is positive, as $\alpha \in (0,1)$,and subtracting $L/(L+K) < 1$ from 2 is > 0. We consider the only case where the second order condition is met such that the denominator becomes negative, in order for $dL/dR_L < 0$. To meet the second-order condition requires equation (3.3) to be negative as shown below:

$$z\alpha(\alpha - 1)(K+L)^{\alpha} + \alpha R_L K < 0 \qquad (3.3)$$

Using equation (3.1) and substituting values for $z\alpha(K+L)^{\alpha-1}$, it is seen that for any value of $\alpha \leq 2/3$ the second-order condition is met. The other comparative static result is dL/dK.

For $dR_L = 0$, $$\frac{dL}{dK} = -\frac{[z(\alpha - 1)Q - R_L L]}{[z(\alpha - 1)Q + R_L K]} \quad (3.4)$$

For $\alpha \leq 2/3$, dL/dK is negative as shown in Appendix III. The special production function mimics the general result. Firms with high collateral will borrow less. However, if the value of collateral slips, then the loan amount will be lower as seen by the mortgage difficulty where home prices are declining in the U.S.

THE THEORETICAL MODEL OF LOAN PROVISION

In this section a simple yet rich model of loan provision is provided to capture the RBC effects on loan issues utilizing the model of Peek and Rosengren (1995). This model incorporates bank equity *E* in the balance sheet constraint which allows banks to maintain different capital requirements for different asset types based on the level of risks.

Thus bank equity *E* helps banks to manage their portfolio-risks by altering as needed. The representative bank faces a simple balance sheet constraint:

$$E + D \equiv L + S \quad (4.1)$$

where loans *L* and securities *S* constitute the assets. Deposits *D*, is assumed to be homogenous and is endogenous. Equity *E*, the net worth of the bank is assumed to be positive. The bank faces a binding capital constraint set by the RBC requirements aimed to prevent large loan losses. It also allows banks to incorporate more than one asset (besides loans) enabling some arbitrage opportunities.

The banks maintain a certain proportion of assets as bank equity:

$$E = aS + bL \quad \text{where} \quad 0 < a < b < 1 \quad (4.2).$$

In equation (4.2), parameters *a* and *b* represent the relative share of assets that must be maintained as a proportion of bank equity *E* (exogenous) prescribed by the regulators for each category of assets *S* and *L*. Higher equity capital allows banks with a cushion to absorb larger loan losses are consistent with the concept of promoting market discipline via recapitalization of banks. The Fed typically chooses $b > a$ as loans become riskier than securities. The numerical values of '*a*' and '*b*' are chosen by the Fed. Anecdotal evidence indicates that during the early '80s regulators

treated loans and securities as equivalent, making *a* equivalent to *b*. During the late 1980s, and in early 2007 loans became increasingly riskier compared to securities and *b* exceeded *a*.

In early 1990s when banks faced new binding capital constraints as a result of loan losses had three alternatives to correct their capital/asset ratio: raise new capital, shrink both assets and liabilities, or reduce the risks of their assets. In practice, a bank's choice to shrink rather than to issue new equity may be due to asymmetric information and the lemon problem, as pointed out by Myers and Majluf (1984). It is a stylized fact in finance literature that only managers of troubled banks have an incentive to issue new equity to cover up for possible loan losses and other problems within their asset portfolios. Potential equity holders realize the predicament and refrain from purchasing any new stocks issued at a lower price, even when the returns are normal. Thus, it is not feasible for the current bank management and share holders to issue new shares. Perhaps the only alternative open to the bank managers would be to shrink both the existing assets and liabilities or to reduce the risks of their assets.

Using the balance sheet constraint (4.1) and the capital constraint (4.2), *S* is eliminated and we solve for *D*. *D* is linearly related to *L* and *E*, as shown in equation (4.3a):

$$D=\frac{(1-a)}{a}E+\frac{(a-b)}{B}L \tag{4.3a}$$

For convenience, equation (4.3a) is simplified to equation (4.3b):

$$D=\alpha(a)E+\beta(a,b)L \tag{4.3b}$$

where $\alpha = (1 - a)/a$ and $\beta = (a - b)/a$ such that $\alpha > 0$ and $\beta < 0$ given $b > a$. To increase *L* for given *E, S* would have to drop by more than *L* rises, as $b > a$. Thus, $S + L$ drops, and, hence, *D* also drops. The interest rate on securities r^S (exogenous) is determined in the securities market. Here commercial banks purchase and sell desired quantities of securities at r^S. So, the bank can choose the optimum amount of securities *S* in the securities market to optimize its portfolio allocation.

In the sprit of the Monti-Klien (1971) imperfectly competitive banking model facing a lending rate r^L for loans and a deposit rate r^D, our model share the same characteristics. We also use inverse functions of r^L and r^D for convenience and they are given as:

$$r^D=f(D) \text{ where } f_D(D)>0 \tag{4.4a}$$

Bank deposits and deposit rates r^D are positively related, whereas lending rates r^L and loan demand are negatively related and are given as:

$$r^L = g(L,K) \text{ where } g_L(L,K)<0,\ g_K(L,K)<0 \tag{4.4b}$$

In equation (4.4b), the derivative $g_K(L,K) < 0$ indicates firms with higher equity will demand fewer loans. Equation (4.4b) has been derived explicitly in the loan demand section.

A PROFIT MAXIMIZING BANK

Banks, like any other firms, are assumed to maximize profits π^B, with respect to the choice of *D, L*, and S, and subject to an adding up constraint and the RBC constraint. The profit function of the bank is given as

$$\pi^B = r^S S + r^L L - \phi(L,K,Y)L - r^D D \tag{5.1}$$

where π^B is the profit function which is the sum of intermediation margins on loans and securities minus banks' managerial expenses and loan loss provisions. Bank revenue comprises total earnings from securities $r^S S$ and total earnings from loan issues $r^L L$. The formulation allows for the expectations of some default fraction $\phi(L, K, Y)$ and interest paid to depositor's $r^D D$. The loan default fraction ϕ are functions of loans *L*, collateral *K*, and gross state product *Y*. We assume $\phi_L > 0$, $\phi_K < 0$, and $\phi_Y > 0$. $\phi_Y < 0$ is negative, implying close association between loan defaults and economic slowdown (Wojnilower, 1980).

Simply manipulating equation (5.1) and substituting leads to express the profit function in terms of the only choice variable *L*

$$\pi^B = r^S[E + \alpha E + \beta L - L] + g(L,K)L - \phi(L,K,Y)L - f(\alpha E + \beta L) \tag{5.2}$$

Differentiating equation (5.2) by *L*, the following first-order condition is obtained:

$$\frac{d\pi^B}{dL} = (\beta - 1)r^S + g_L(L,K) + g(L,K) - \phi_l(L,K,Y)L - \phi(.) - f'(.)\beta D - \beta f(D) = 0 \tag{5.3}$$

Differentiating equation (5.3) twice, the second-order condition is obtained:

$$\begin{aligned} \frac{d^2\pi^B}{dL^2} = {} & g_{LL}(L,K)L + 2\,g_L(L,K) - \phi_{LL}(L,K,Y)L \\ & - 2\phi_L(.) - 2f'(.)\beta^2 - f''(.)\beta^2 D \end{aligned} \tag{5.4}$$

We make restrictive assumptions for $d^2\pi^B/dL^2 < 0$ to hold and they are as follows: $\phi_{LL} \geq 0$, $g_{LL} \leq 0$, and $f''(.) \geq 0$ (the demand curve is strictly convex) are sufficient to guarantee that the second order condition holds provided g_{LL} is not large and positive. The bank will set forth the loan amount to the point where the marginal amount of loans is equated to the opportunity costs of the returns on loans, and other funds, and the state of the economy as well as the collateral.

COMPARATIVE STATIC RESULTS FOR THE EQUILIBRIUM LOAN QUANTITY *L*

In the light of new capital requirements, an increase in equity capital will allow banks to further their lending assuming that $d^2\pi^B/dL^2 < 0$. For deriving any comparative result (e.g., dL/dE), the relevant second order condition is used. The necessary condition for comparative static is:

$$\frac{\partial^2\pi^B}{\partial L\partial E}\, dE + \frac{\partial^2\pi^B}{\partial L^2}\, dL = 0 \tag{6.1}$$

Using (6.1) and switching sides, we get the expression for dL/dE:

$$\frac{dL}{dE} = -\frac{\frac{\partial^2\pi^B}{\partial L\partial E}}{\frac{\partial^2\pi^B}{\partial L^2}} = \frac{\alpha\beta f''(.)\, D + 2f'(.)\,\beta\alpha}{\frac{\partial^2\pi^B}{\partial L^2}} > 0 \tag{6.2}$$

Using previous assumptions that $f'' \geq 0$ and $f' > 0$, and recalling that $\alpha > 0$ and $\beta < 0$, dL/dE is > 0 suggests that an increase in equity capital makes it possible for banking firms to provide additional loans to their loan customers. Thus, an increase in the bank equity will push banks to increase the total loan volume provided the new capital standards are met.

The other comparative static results that requires attention are the impact of the regulatory parameters *a* (capital requirement against securities) and *b* (capital requirement against loans) on equilibrium lending *L*. The effect of *a* on *L* is given as:

$$\frac{dL}{da} = -\frac{\frac{\partial^2\pi^B}{\partial L\partial a}}{\frac{\partial^2\pi^B}{\partial L}} \quad \text{where} \quad \frac{\partial^2\pi^B}{\partial L\partial a} = \frac{b}{a^2}[r^S - r^D(1+\varepsilon_{r_D,D}) - \beta D f''(.)\, L - 2\beta f'(.)\, L] \tag{6.3}$$

$$-\frac{1}{a^2}[-2\beta f'(.)\, E - f''(.)\,\beta DE] \text{ and } r^S > r_D(1+\varepsilon_{r_D,D}) \text{ holds}$$

general, the sign of equation (6.3) is negative because an increase in *a* has opposite effects: For given *E*,

(i) an increase in *a* will require the bank to shrink assets, and that can occur when *L* declines.
(ii) an increase in *a* will require the substitution between assets *L* and *S*. Since *S* is subject to higher capital requirements, *L* will be substituted for *S*, thereby increasing *L*.

Similarly, an increase in *b* gives:

$$\frac{dL}{db} = -\frac{\frac{\partial^2 \pi^B}{\partial L \partial b}}{\frac{\partial^2 \pi^B}{\partial L^2}} < 0 \quad \text{where } \frac{\partial^2 \pi^B}{\partial L \partial b} = \beta_b \left[(r^S - r^D (1 + \varepsilon_{r_D, D})) - \beta L (f''(.) LD + 2f'(.)) \right] < 0$$
$$\text{provided } r^S > r^D (1 + \varepsilon_{r_D, D}) \tag{6.4}$$

If the restriction mentioned above holds, then *dL/db* is negative. An increase in *b* would cause disincentive for banks to extend additional loans for two simple reasons: For given E,

(i) an increase in *b* will require banks to shrink assets and that can happen when *L* decreases.
(ii) an increase in *b* will require banks to substitute away from *L* and into *S* lessening *L*.

How banks react to changes in collateral K (firms' collateral) on L is given as

$$\frac{dL}{dK} = -\frac{\frac{\partial^2 \pi^B}{\partial L \partial K}}{\frac{\partial^2 \pi^B}{\partial L^2}} = -\frac{[L g_{LK}(.) K + g_K(.) - \phi_{KL}(.) L - \phi_K(.)]}{\frac{\partial^2 \pi^B}{\partial L}} > 0 \text{ or } < 0 \tag{6.5}$$

The effect of a change in K on L is also ambiguous. First, increased *K* will permit banks to lend more; L is likely to increase. Secondly, increased *K* will allow borrowers to use retained earnings and reduce loan demand. Therefore, the net effect of *K* on *L* is unclear. However, in the context of the current mortgage crisis we expect *dL/dK* to be negative. As home prices have softened, so is the home equity lessening *L*. The effect of state gross product Y on equilibrium loan quantity is illustrated below. As one would expect, dL/dY should be positive, where Y captures the condition of the economy. The comparative static for dL/dY is:

$$\frac{dL}{dY} = \frac{[\varphi_{LY}(.)L + \varphi_{Y}(.)]}{\frac{d^2\pi^B}{dL^2}} > or < 0 \qquad (6.6)$$

dL/dY is positive, provided $\phi_{LY} \leq 0$. As economy slows down so will be the equilibrium loan amount. According to Wojnilower (1980), during business cycle peaks, credit demands are virtually interest inelastic and credit supply falls way below credit demand, causing a paralytic effect on the economy. In 2007 September, Ben Bernanke said that the downturn in the residential construction has trimmed the annual rate of U.S. economic growth about ¾ percentage point on average over the past year and a half. Thus in this model $dL/dY < 0$ due to economic slowdown.

The sign for dL/dr^S should be negative; an increase in the alternative rates tends flow of funds in that direction because of arbitraging opportunities. The effect is:

$$\frac{dL}{dr^S} = -\frac{\frac{\partial^2\pi^B}{\partial L \partial r^S}}{\frac{\partial^2\pi^B}{\partial L^2}} = -\frac{(\beta - 1)}{\frac{\partial^2\pi^B}{\partial L}} < 0 \quad \text{where} \beta < 0 \qquad (6.7)$$

So, in the loan disbursement model, loan quantity is a function of *E, a, b, K, Y,* and r^S which will form the basis for the empirical work.

COMPARATIVE STATIC RESULTS OF THE EQUILIBRIUM LOAN RATE R L

In order to derive the comparative static results of the lending rate r^L, the following information in equation (4.4b) is used. Totally differentiating equation (4.4b) and setting *dK* equal to zero results in the following expression:

$$dr^L = g_L dL \quad where g_L < 0 \qquad (7.1)$$

Dividing dr^L (equation 7.1)) by *dE*, *da*, *db*, dr^S, and *dY* respectively and making use of the results obtained from equations (6.2) - (6.7), except for equation (6.5), will provide us all the comparative static results needed for r^L. The only exception will be dr^L/dK, which comes directly by differentiating equation (4.4b). Since expressions *dL/da* and *dL/dK* are ambiguous, this implies signs of dr^L/da and dr^L/dK are all ambiguous.

The effect of *E* on r^L is as follows:

$$\frac{dr^L}{dE} = g_L \frac{dL}{dE} \tag{7.2}$$

In equation (6.2), dL/dE is > 0, and g_L is < 0 in equation (4.4b) making $dr^L/dE < 0$. As E goes up so does the quantity of loans for given S, and high loan quantity can only occur when r^L is low.

$$\frac{dr^L}{db} = g_L \frac{dL}{db} \tag{7.3}$$

However, dL/db is < 0, making dr^L/db positive. Intuitively speaking, as loans become more risky and banks have to hold more equity, they will have a tendency to charge higher interest rates on new loans. Next, consider the effect of Y on r^L where:

$$\frac{dr^L}{dY} = g_L \frac{dL}{dY} < 0 \tag{7.4}$$

It seems reasonable to say that as Y increases, the marginal cost curve of loans shifts to the right making r^L go down, making $dr^L/dY<0$. This argument can be extended to explain the current credit crunch as well where economic weakness is causing interest rates to rise as banks are skeptical about the credit worthiness of their borrowers and use higher rates to reduce moral hazard problem. It is a stylized fact in finance and economics that interest rates move in tandem as shown in equation (7.5):

$$\frac{dr^L}{dr^S} = g_L \frac{dL}{d\,r^S} > 0. \tag{7.5}$$

CONCLUSION

This article examines the determinants of the credit crisis in the U.S. and its effect on the economy. The model enables us to identify two things with regards to the current credit crunch. One, as house prices have softened, so is the home equity (K). Thus, less K implies fewer loans available for marginal borrowers. Banks responded by tightening lending standards and made it more difficult to obtain loans thereby adding to the current credit bottleneck. Two, the rise in recent mortgage rates along with the switch from short term lower fixed rates to higher adjustable rates has seen a sharp increase in delinquencies especially for the sub-prime market. Thus, the credit problem

has affected a much broader market beyond the home builders industry exposing the weaknesses of U.S. economy. The model explains the 2007 credit crisis at least in part. From the theoretical view point it seems quite plausible that the current crunch is caused by the "supply" side factors.

This paper also addresses the issue of moral hazard by showing the importance of collateral in a loan market. Using a generalized production function, and a binding capital constraint, the preliminary results indicate that the influence of collateral on loan amount and lending rates on loan quantity are ambiguous. The ambiguity stems from the presence of limited liability ($p_0 \geq 0$) where banks realize that there will be no loan repayment if price dips below p_0. For the specific production function, however, the results are conclusive provided the parametric restriction, $\alpha \leq \frac{2}{3}$ (strictly) holds. Clearly, the model supports the regulatory shift towards recapitalization of the depository institutions introduced in early 1990s which is well received in the banking literature because of its success, where the new capital guidelines has prevented serious bank runs numbering only 120 since 1992. These developments in the banking regulation are consistent with the public policy measures. However, I believe, the RBC requirements directly impact the lending by skewing the market, lowering market efficiency, and, raising the capital cost to the marginal customers – imminent result the credit crunch.

ENDNOTES

1 In the literature credit crunch is typically defined in several ways. The first way of defining credit crunch is the leftward shift in the supply of bank credit for the existing loan demand (Berger and Udell 1994). The second is a reduction in the equilibrium loan quantity either due to a decline in loan demand, or a decline in loan supply, or both (Bernanke and Lown 1991). The third is associated with cyclical components of the business cycle which defines significant retardations or reductions in credit and aggregate demand occurs only when there is an interruption in the supply of credit – "a credit crunch" (Wojlilower 1980). The fourth is "the widespread reluctance of banks to lend to borrowers to whom they would have lent eagerly in the past," which had broad effects on the U.S. economy (Wessel, 1991).

2 The U.S flow of funds data show that in 1990, banks extended only $2 billion net new loans to non-financial corporations compared to $33 billion new net loans in both 1988 and 1989.

3 I am going to use the assumption of concavity as a sufficient condition to achieve the first [$Q'(.) > 0$] and second order [$Q(.)'' < 0$] conditions of profit maximization when using the generalized production function.

4 dp_0/dL depends on the elasticity of demand for the quantity of production with respect to loans where $dp_0/dL = R_L(1- \epsilon_{L, Q(.)})/ Q$ and ϵ = percentage change in production/percentage change in loan. If $\epsilon_{L,Q(..)}$ is inelastic, then $dp_0/dL > 0$. If the firm in this study is operating in a competitive environment and produces at a point where the average product of input is maximized (the slope of AP = 0) and MP = AP, this implies that at competitive equilibrium, $dp_0/dL = 0$ when AP = MP and $\epsilon = 1$. Alternatively using the concept of limited liability on the random variable p lies between z and p_0. Any value of $p \leq p_0$ will theoretically make $dp_0/dL = 0$ because of restrictions on the price level imposed by limited liability, and this condition is itself sufficient for the second-order condition to hold.

[5] The detailed derivation is shown in appendix 1.

REFERENCES

Avery, R.B., and Berger, A. N., 1991. Risk-Based Capital and Deposit Insurance Reform. *Journal of Banking and Finance* 15, 847-74.

Baltesnsperger, E., 1980. Alternative Approaches to the theory of Banking Firm. *Journal of Monetary, Economics* 102,1-37.

Benston, G. J.,1991. Is There a Credit Crunch and What Should We Do About It? *Journal of Applied Corporate Finance* 41, 86-91.

Berger, A. N., and Udell G. F., 1992. Some evidence of the Empirical Significance of Credit Rationing. *Journal of Political Economy* 100, 1047-77.

Berger, A. N., and Udell G. F., 1994. Did Risk-Based Capital Allocate Bank Credit and Cause a Credit Crunch in the United States. *Journal of Money, Credit, and Banking*, 263,583-628.

Berger, A. N., and Udell G. F., 1995. Relationship Lending and Lines of Credit in Small Firm Finance. *Journal of Business,* 68 3,351-63.

Bester, H. 1985. Screening vs Rationing in Credit Markets with Imperfect Information. *American Economic Review* 754,850-55.

Bernanke, B. 2007 (Kansas City Fed's Jackson Hole, WY Conference August 31,2007)

Bernanke B S., and Lown, C.S. 1991.The Credit Crunch. *Brookings Paper on Economic Activity*, 205-47.

Bizer, David S., 1993. "Regulatory Discretion and the Credit Crunch." *Working Paper*, U.S. Securities and Exchange Commission, Washington, D.C.

Blinder, A. S., Jun. 1987. Credit Rationing and Effective Supply Failures. *Economic Journal*, 327-55.

Credit Crunch of 1990. *Journal of Applied Corporate Finance*, 41, 92-95.

De Meza, D., and Webb, D., 1987. Too Much Investment, A Problem of Asymmetric Information. *Quarterly Journal of Economics* 1022, 281-92.

Diamond, D., 1991. Monitoring and Reputation, The Choice between Bank Loans and Directly Placed Debt. *Journal of Political Economy* 99, 689-721.

Flannery M., 1991. Pricing Deposit Insurance When the Insurer Measures Bad Risk With Error. *Journal of Banking and Finance* 15, 975-998.

Freixas X., and Rochet J.C., 1997. Microeconomics of Banking, The MIT Press.

Gennotte, G., and Pyle, D., 1991.Capital Controls and Bank Risk. *Journal of Banking and Finance* 15, 805-824.

Gertler, M., and Hubbard, R. G., 1993 Corporate Financial Policy, Taxation, and Macroeconomic Risk. *Rand Journal of Economics,* Summer, 286-303.

Hancock, D., and Wilcox, J.A.,1992. The Effects of Bank Assets of Business Conditions and Capital Shortfalls, In Credit Markets in Transition. *Proceedings of the 28th Annual Conference on Bank Structure and Competition*, Federal Reserve Bank of Chicago, 502 -20.

Haulbrich, J.G., and Watchel, P., 1993. Capital Requirements of Shifts in Commercial Bank Portfolios. *Federal Reserve Bank of Cleveland Economic Review* 29, 2-15

Holmström, B., and Tirole, J., 1993. Financial Intermediation, Loanable Funds, and the Real Sector.IDEI, Toulouse University. Mimeograph.

Keeley, M.C., and Furlong, F.T., 1990. A Reexamination for Mean-variance Analysis of Bank Capital Regulation. Journal of Banking and Finance 14, 69-84.

Klein, L. A.,1971 A Theory of Banking Firm. *Journal of Money Credit and Banking*, 205-18.

Kim, D and Santomero, A.M. 1988. Risk in Banking and Capital Regulation. *Journal of Finance* 43, 1219-1233.

King, S.R., 1986. Monetary Transmission, through Bank Loans or Bank Liabilities? *Journal of Money, Credit, and Banking* 18, 290-303.

Melitz, J,, and Pardue, M., 1973. The Demand and Supply of Commercial Bank Loans. *Journal of Money Credit and Banking*, 669-691.

Monti, M., 1972 *Deposits, Credit, and Interest Rate Determination Under Alternative Bank Objectives. In Mathematical Methods in Investment and Finance*, edited by G.P. Szego and K. Shell. Amsterdam, North Holland.

Mussa, M., and Rosen, S., 1978. Monopoly and Product Quality. *Journal of Economic Theory* 182, 301-17

Myers, S.C., and Majluf. N.S., 1984. Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not. *Journal of Financial Economics*,184-221.

Peek, J. and Rosengren. E., 1995. The Capital Crunch, Neither a Borrower nor a Lender Be. *Journal of Money Credit, and Banking* 273,625-638.

Shrieves, R.E. and Dahl, D.,1993. Regulation, Recession, and Bank Lending Behavior, The 1990 Credit Crunch". University of Tennessee *Working Paper*.

Stiglitz, J.E, and Weiss, A., 1981. Credit Rationing in Markets with Imperfect Information. *American Economic Review* 713, 393-410.

Syron, Richard F., 1991. Are We Experiencing a Credit Crunch? Federal Reserve Bank of Boston, *New England Economic Review*, 3-10.

Thakor, A.V., 1993 Capital Requirements, Monetary Policy, and Aggregate Bank Lending. Indiana University *Working Paper*.

Waite, S.R., 1991.The Eclipse of Growth Capital. *Journal of Applied Corporate Finance*, 41, 77-85.

Wojnilower, A. M., 1980. The Central Role of Credit Crunches in Recent Financial History. *Brookings Paper on Economic Activity.*2, 277-326.

Yago, G. 1991. The Credit Crunch A Regulator Squeeze on Growth Capital. *Journal of Applied Corporate Finance*, 41, 96-100.

APPENDIX 1

Differentiating (3.1) with respect to L and using Leibnitz's Rule results in the following:

$$\frac{dE(\pi^F)}{dL} = \int_{p_0}^{z} [\,\alpha(K+L)^{\alpha-1}\, p - R_L]\frac{1}{z}dp + \frac{dz}{dL}(zQ(.) - R_L L)\frac{1}{z} - \frac{dp_0}{dL}(p_0 Q - R_L L)\frac{1}{z} = 0$$

In the above equation, the last two terms on the right hand side drops, as z is constant and by plugging $p_0 = R_L L/(K+L)^{\alpha}$ into the equation. Thus, the first order condition becomes

$$\int_{p_0}^{z} [p\,\alpha((K+L)^{\alpha-1} - R_L]dp = 0 \qquad \text{(A1.2)}$$

To check for the second order condition for profit maximization with the specific production function, consider the following:

$$\frac{d^2E(\pi^F)}{dL^2} = \int_{p_0}^{z} p\alpha(\alpha-1)(K+L)^{\alpha-2}\frac{1}{z}dp - \frac{dp_0}{dL}(p_0\,\alpha(K+L)^{\alpha-1} - R_L)\frac{1}{z} < 0 \qquad \text{(A1.3)}$$

For $d^2E^2(\pi^F)/dL^2 < 0$ in equation (3.4), we signed the term associated with dp_0/dL and proved that the profit function is concave.

Next, it is necessary to define for the uniform distribution, $f(p) = 1/z$, where $p = [0, z]$, the following:

$$1 - F(p_0) = \int_{p_0}^{z} f(p)dp = \frac{1}{z}\int_{p_0}^{z} dp = \frac{1}{z}(z - p_0) \qquad \text{(A1.4)}$$

$$G(p_0) = \int_{p_0}^{z} pf(p)\,dp = \frac{1}{z}\int_{p_0}^{z} p\,dp = \frac{1}{2z}(z^2 - p_0^2) \qquad \text{(A1.5)}$$

$$H(p_0)=\frac{G(p_0)}{1-F(p_0)}=\frac{1}{2}(z+p_0) \quad \text{where } p_0=\frac{R_L L}{(K+L)^{\alpha}} \tag{A1.6}$$

We note that $dH/dp_o = ½$ for the uniform distribution by differentiating equation (3.7) with respect to p_0. Using equation (3.7), the first order condition is rewritten as

$$z\alpha(K+L)^{\alpha-1}=R_L[2-\alpha\frac{L}{(K+L)}] \tag{A1.7}$$

Totally differentiating equation (A1.7) results in equation (A1.8).

$$[z\alpha(\alpha-1)(K+L)^{\alpha-2}+\alpha R_L\frac{K}{(K+L)^2}]dL+[z\alpha(\alpha-1)(K+L)^{\alpha-2} - \alpha R_L\frac{K}{(K+L)^2}]dK=[2-\frac{L}{(K+L)}]dR_L \tag{A1.8}$$

Using equation (A1.8), I determine the signs for dL/dR_L and dL/dK, respectively. I expect $dL/dR_L < 0$, under the restrictive condition where $\alpha \leq ⅔$. Rearranging the terms in equation (A1.8) and changing sides gives the expression for dL/dR_L and dL/dK given by equation (3.2), and equation (3.4).

APPENDIX 2

Here I will show that the expected profit is concave, a sufficient condition for profit maximization ($d^2E(\pi^F)/dL^2 < 0$). We make use of the equations (3.3), (3.4), and (3.6), and the relation dp_0/dL for our purpose. Using equation (3.1), and substituting $Q = (K+L)^{\alpha}$ the new p_0 becomes $p_0 = R_L L/(K+L)^{\alpha}$ where $\alpha \in (0,1)$. In equation (3.4) we need to demonstrate that dp_0/dL and the term associated with it

$$\frac{d^2E(\pi^F)}{dL^2}=\int_{p_0}^{z} p\alpha(\alpha-1)(K+L)^{\alpha-2}\frac{1}{z}dp-\frac{dp_0}{dL}(p_0\alpha(K+L)^{\alpha-1}-R_L)\frac{1}{z}$$

$$\text{where } \frac{dp_0}{dL}=\frac{R_L}{Q}[1-L\frac{\alpha}{(L+K)}]$$

$$G(p_0)=\int_{p_0}^{z} pf(p)dp=\frac{1}{z}\int_{p_0}^{z} pdp=\frac{1}{2z}(z^2-p_0^2)$$

is non-negative. It will be adequate to establish that the expected profit function is concave.
I use equation (3.6) and $G(p_0)$ extensively for the analysis. Substituting values of $G(p_0)$ and p_0 in $d^2E(\pi^F)/dL^2$ and simplifying, we have

$$\frac{d^2E(\pi^F)}{dL^2} = \frac{\alpha(\alpha-1)Q}{2(K+L)^2}[z^2 - p_0^2] + \frac{p_0^2 Q}{L^2}\left(1 - \alpha\frac{L}{K+L}\right) \qquad \text{(AII.1)}$$

$$\text{where} \quad z^2 = p_0^2\left[\frac{2(K+L)}{\alpha L} - 1\right]^2$$

where z^2 in (A1) is derived by using equation (3.8), $p_0 = R_L L/Q$, $G(p_0)$ and squaring.

Substituting z^2 in $[z^2 - p^2]$ in the above relation, we get:

$$\frac{d^2E(\pi^F)}{dL^2} = \frac{\alpha(\alpha-1)Q}{2(K+L)^2}\left[p_0^2\left(\frac{2(K+L)}{\alpha L} - 1\right)^2 - p_0^2\right] + \frac{p_0^2 Q}{L^2}2\left(\alpha\frac{L}{K+L} - 1\right) \qquad \text{(AII.2)}$$

Eliminating Q and p_0 from (A2) and rearranging terms, we have (A3):

$$\frac{d^2E(\pi^F)}{dL^2} = \frac{\alpha(\alpha-1)}{2(K+L)^2}\left[\left(\frac{2(K+L)}{\alpha L} - 1\right)^2 - 1\right] + \left(\frac{\alpha L}{K+L} - 1\right)\left(\frac{1}{L^2}\right) \qquad \text{(AII.3)}$$

Further simplifying (A3) and canceling terms, we finally have

$$\frac{d^2E(\pi^F)}{dL^2} \equiv \frac{(\alpha-1)}{L(K+L)}\left[\frac{K+L}{\alpha L} - 1\right] + \frac{Q}{L^2}\left(\frac{\alpha L}{K+L} - 1\right) \qquad \text{(AII.4)}$$

For the second order condition to hold, $d^2E(\pi^F)/dL^2$ should be negative. This is true since $(\alpha-1) < 0$ and $(\alpha L/(K+L) - 1)$ is also negative.

APPENDIX 3

We need to show that for $dL/dR_L < 0$ and $dL/dK < 0$ to be true, under the indicated production function it is sufficient that $\alpha \leq 2/3$.

$$H(p_0) = \frac{G(p_0)}{1 - F(p_0)} = \frac{1}{2}(z + p_0) \text{ where } p_0 = \frac{R_L L}{(K+L)^{\alpha}} \qquad \text{(AIII.1)}$$

$$\text{and } z\alpha(K+L)^{\alpha-1} = R_L\left[2 - \alpha\frac{L}{(K+L)}\right]$$

We need to show that the expression in equation (3.3) is negative:

$$z\alpha(\alpha-1)(K+L)^{\alpha} + \alpha R_L K < 0$$

Using the relation $z\alpha(\alpha-1)(K+L)^{\alpha-1} = R_L[2 - \alpha(L/(K+L))]$ and substituting in equation (A3.III) and eliminating R_L, we have

$$(\alpha - 1)[2(K+L) - \alpha L] + \alpha K < 0$$
$$\text{or, } 2\alpha L + 2\alpha K - 2L - 2K - \alpha^2 L + \alpha L + \alpha K < 0$$
$$\text{or, } 3(\alpha L + \alpha K) < 2(L+K) + \alpha^2 L$$
$$\text{or, } (3\alpha - 2)(L+K) - \alpha^2 L < 0$$
$$\text{if } 3\alpha - 2 \leq 0 \text{ i.e., } \quad \alpha \leq \frac{2}{3}$$

Thus, $\alpha \leq 2/3$ reiterates that the production function exhibits diminishing marginal returns and is concave: a condition that is sufficient for profit maximization to hold

This is a combined edition containing both Volume 9, Number 1, and Volume 9, Number 2

Articles for Volume 9, Number 2

AN ANALYSIS OF ALTERNATIVE METHODOLOGIES AND INTERPRETATIONS OF MORTGAGE DISCRIMINATION RESEARCH USING SIMULATED DATA

Christopher L. Brown, Western Kentucky University
W. Gary Simpson, Oklahoma State University

ABSTRACT

Over the years, there has been intense debate on the significance of default rates in identifying or ruling out discriminatory mortgage lending and on the appropriate methodology to use in testing for mortgage discrimination. This paper reviews the debate and uses simulated data to provide conclusive evidence on the merits of the alternative theories and methodologies. Understanding the relationship of default rates as potential indicators of discrimination and assuring that the methodology used in these studies is appropriate is very important because the findings of mortgage discrimination research may be used to influence public policy.

The findings indicate that we cannot focus on default analysis to answer the question of whether or not discrimination is occurring. The findings also indicate a well-specified logit regression model may be used to identify mortgage discrimination that occurs after completion of the loan application.

INTRODUCTION

For the past several years, researchers have focused on investigating discrimination in the residential mortgage market. New data and methodologies have been employed and new theoretical lending models have been developed to explain why loan approval rates are higher for non-minority applicants than for minority applicants. There has been intense debate on the significance of default rates in identifying or ruling out discriminatory lending and on the appropriate methodology to use in testing for discrimination. This paper reviews the debate and uses simulated data to provide conclusive evidence on the merits of the alternative theories and methodologies. Understanding the relationship of default rates as potential indicators of discrimination and assuring that the methodology used in these studies is appropriate is very important because the findings of discrimination studies may be used to influence public policy.

The mortgage discrimination debate intensified with the release of a study conducted by researchers at the Federal Reserve Bank of Boston (Munnell, Tootell, Browne and McEneaney

(1996)). The study employs the most comprehensive loan application information of any of the recent discrimination studies. The authors use a logit regression equation that includes all of the variables that should be relevant to the loan decision. The race of the applicant is included as an additional explanatory variable. If the coefficient on race is significant, it is interpreted as evidence of discrimination.

Munnell *et. al.* (1996) find minority applicants, on average, have greater debt burdens, higher loan-to-value ratios, and weaker credit histories than non-minorities. Furthermore, denied minorities have lower income and wealth, higher obligation and loan-to-value ratios, and worse credit histories than denied non-minorities. Despite these facts, the authors find evidence of discrimination against minorities. They find that minority applicants are rejected 60 percent more often than non-minority applicants when financial, employment, and neighborhood characteristics are held constant.

The findings of Munnell *et. al.* (1996) have received a great deal of attention from policymakers and academic researchers. Most of the debate focuses on perceived shortcomings in the study. Criticisms of the Munnell *et. al.* (1996) study include problems with the integrity of the data (Horne (1994), Liebowitz (1993), and Carr and Megbolugbe (1993)), the authors' failure to consider default rates (Becker (1993), England (1993), and Brimelow and Spencer (1993)), and problems with the use of a single-equation logit regression model of the probability of loan approval to detect discrimination.

More recently, Blank et al (2005) investigated racial discrimination in mortgage lending in Washington, DC. Using three different methodologies, a dissimilarity index approach, a three-way crosstabulation approach and a logistic regression. The adjusted dissimilarity index approach is based on the theory that, after considering for differences in neighborhood factors and using variables on the loan applicants that are available through HMDA, approval rates should be approximately the same across census tracts. Blank et al (2005) find there is a disparity between census tracts. They conclude that 10.64 percent of loans that should have gone to underserved census tracts were denied. That amounts to 1,315 loans. The crosstabulation approach simply evaluates whether there is a disparity in lending based on only income and race. After considering income, they find a significant difference in the proportion of loans denied between minorities and non-minorities across all income levels. The third approach is a logistic regression model. The model includes the race of the applicant along with neighborhood characteristics. They again find that minorities are less likely to receive loans, after accounting for the factors in the model.

The findings by Blank et al (2005) do not necessarily indicate discrimination. None of the methodologies employed by Blank et al (2005) include the credit history of the applicants. As explained in more detail in Section II, the distribution of credit quality may explain differences in loan approval rates even after considering all of the variables used in the Blank et al (2005) study.

Section II discusses the role of default rates in interpreting the results of mortgage discrimination research. A model of the relationship between loan denial and default rates and discrimination developed by Ferguson and Peters (1995) is presented. Section III discusses the criticisms of the use of a single-equation logit regression equation to measure mortgage

discrimination. Section III also presents the reverse regression methodology LaCour-Little (1996) applies to the Munnell *et. al.* (1996) data to test for mortgage discrimination. Weaknesses in the reverse regression methodology are also discussed in Section III. Section IV presents the simulation analysis used to show the relationship between default rates and discrimination and to compare the performance of a single-equation logit regression model and the reverse regression model in identifying discriminatory lending. The simulation results are presented in Section V. The conclusion and recommendations for future research are presented in Section VI.

THE ROLE OF DEFAULT RATES IN MORTGAGE DISCRIMINATION RESEARCH

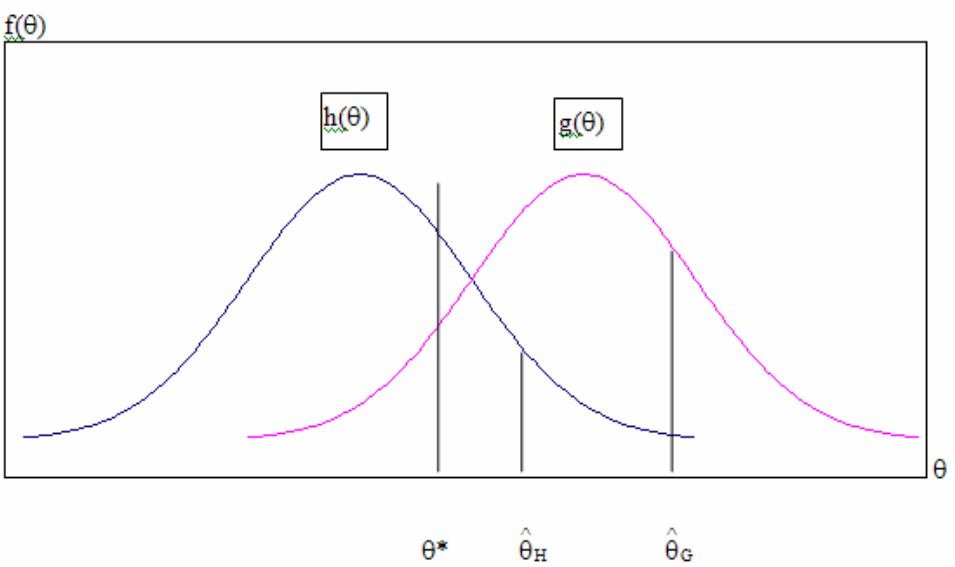
Becker (1993) and England (1993) argue that, if discrimination exists, minorities should have lower default rates than non-minorities. They contend that failing to observe lower default rates for minority borrowers is evidence against racial discrimination in mortgage lending. Brimelow and Spencer (1993) also use this reasoning to challenge the findings of Munnell *et. al.* (1996). They cite the Boston Fed's finding that the average default rate for minority neighborhoods in Boston is the same as the rate for non-minority neighborhoods. Brimelow and Spencer (1993) argue that equal default rates for minority and non-minority neighborhoods contradicts the Munnell *et. al.* (1996) conclusion that Boston area lenders discriminate against minority applicants.

Munnell *et. al.* (1996), Tootell (1996), Browne and Tootell (1995), Galster (1993), and Ferguson and Peters (1995) argue that racial discrimination in the mortgage market will result in lower default rates for minority borrowers only if certain conditions hold. Tootell (1993) and Browne and Tootell (1995) argue equal minority and non-minority default rates can only be used as evidence of discrimination if the distribution of the quality of accepted minority applicants is identical to the distribution of accepted non-minority applicants.

Ferguson and Peters (1995) explain the relationship between denial rates, default rates, and the distribution of credit quality. They present a model where the distribution of credit quality is higher for non-minority applicants than for minority applicants. This is referred to as heterogeneous credit quality. Recent empirical evidence tends to support the hypothesis that the distribution of credit quality is heterogeneous.

Let q represent the probability of loan repayment. The distributions of credit quality for minority applicants, h(q), and for non-minority applicants, g(q), are shown in Table 1. For simplicity, assume q is measured without error. All applicants with q above some arbitrary cutoff point, q* are approved and applicants with q below q* are denied. All applicants face the same cutoff point, but the average credit score for approved minority applicants, q_h, is lower than the average credit score for approved non-minority applicants, q_g.

Table 1: Ferguson and Peters (1995) Model of Heterogeneous Credit Quality



The Ferguson and Peters (1995) model predicts that approved non-minority applicants will, on average, have higher credit quality than approved minority applicants. The Ferguson and Peters (1995) model indicates that if minorities have a lower distribution of credit quality than non-minorities, they may have higher default rates even if discrimination is present.

MODEL SPECIFICATION ISSUES IN MORTGAGE DISCRIMINATION RESEARCH

Rachlis and Yezer (1993) and Yezer, Phillips, and Trost (1994) argue that single-equation models cannot be used to test for discrimination because of the complexity of the mortgage lending process. The single-equation models do not take into consideration the borrower's choice of loan terms or the borrower's default decision. Yezer, Phillips, and Trost (1994) show the effects of simultaneity and self-selection bias that result from using a single-equation model of the loan approval decision to detect mortgage discrimination. They show that the coefficient on the discrimination variable will be biased upwards.

LaCour-Little (1996) also criticizes the methodology used by Munnell *et. al.* (1996). He also contends that the direct logit regression model produces a biased estimate of the discrimination coefficient. However, his line of reasoning is different than Yezer, Phillips, and Trost (1994). LaCour-Little (1996) implies that the logit regression methodology is inappropriate when one group of applicants has a lower distribution of credit quality than the other group. He states that

differences in average credit quality require the use of a different methodology. He proposes reverse regression as an alternative methodology to detect discrimination in mortgage lending.

LaCour-Little (1996) uses reverse regression on the data from Munnell *et. al.* (1996) to test for discrimination. First, he estimates a direct logit regression equation with eleven independent variables used in Munnell *et. al.* (1996). The dependent variable, ACTION, equals 1 if the loan was denied. There is no race coefficient in this model. The coefficients generated from the regression are used to estimate the probability of loan denial for each observation. The predicted probabilities are considered the inverse qualifications index, Q-INDEX. The Q-INDEX variable is a measurement of the probability of loan denial, therefore higher values of Q-INDEX are bad.

The Q-INDEX values are used as the dependent variable in the following ordinary least squares regression:

$$Q\text{-}INDEX = b_0 + b_1 ACTION + b_2 RACE = e$$

where ACTION equals 1 if the loan was denied and 0 if the loan was approved, and RACE equals 1 if the applicant is a minority and 0 if the applicant is a non-minority. LaCour-Little (1996) contends the coefficient on RACE measures "the excess probability of default required to turn down a minority applicant." LaCour-Little (1996) also calculates a value, a*, that is a measure of the average qualifications of accepted minority applicants relative to accepted non-minority applicants. This measure is calculated as shown:

$$a^* = -b_2/b_1.$$

The results of the LaCour-Little (1996) reverse regression on the loan application data are shown in Table 2. LaCour-Little (1996) concludes the a* value of -.193 indicates that accepted minority applicants had average qualifications 19 percent lower than accepted white applicants. He interprets the RACE coefficient of .057 as the excess probability of default required to reject a minority applicant. He concludes that lenders appear to apply less stringent underwriting standards to minority loan applications, and that there is evidence of reverse discrimination.

Table 2: LaCour-Little (1996) Reverse Regression Results

Variable	Parameter Estimate
Intercept	.0892
ACTION	.296
RACE	.057
a* = -.057/.296 = -.196	

The statistical analysis conducted by LaCour-Little (1996) may be accurate, but the conclusions derived from the analysis are questionable. LaCour-Little (1996) uses the finding that accepted minority applicants have average qualifications 19 percent lower than accepted non-minority applicants to conclude that there is evidence of reverse discrimination. However, the finding that accepted minorities have lower average qualifications than accepted non-minority applicants is not evidence of discrimination. This is the expected outcome based on the Ferguson and Peters (1995) model.

The correct interpretation of the reverse regression methodology employed by LaCour-Little (1996) is relatively straightforward. Since the ACTION and RACE variables are both binary, there are only four possible outcomes calculated by the model. Those four outcomes represent the average Q-INDEX values for (1) approved non-minority applicants (ACTION=0, RACE=0), (2) approved minority applicants (ACTION=0, RACE=1), (3) denied non-minority applicants (ACTION=1, RACE=0), and (4) denied minority applicants (ACTION=1, RACE=1).

LaCour-Little (1996) finds that approved white applicants have an average Q-INDEX of .0892 while approved minority applicants have a Q-INDEX of .1462. These results are consistent with the Ferguson and Peters (1995) model and the findings by Munnell *et. al.* (1996) that minority applicants, on average, have lower income levels, higher debt ratios, and worse credit histories than non-minority applicants. The findings do not indicate reverse discrimination or the absence of discrimination. The only way to prove discrimination (or reverse discrimination) is to use a model that will determine if the marginal cutoff point is the same for minority and non-minority applicants. When the distribution of credit quality is heterogeneous, calculations of differences in average credit quality do not provide useful information about the presence or absence of discrimination.

SIMULATION ANALYSIS

Simulation analysis is conducted to identify the role of default rates in identifying discrimination and to compare the performance of the reverse regression model and the logit regression model in detecting discrimination when non-minority and minority applicants have heterogeneous credit quality. Assume (following Ferguson and Peters (1995)) the screening process leads to a single credit score. The credit score reflects the probability of repayment, q. A uniform, nondiscriminatory credit policy requires that all applicants who meet some minimum required credit score, q*, be approved.

Assume the lender collects all relevant information and inputs it into a logit regression equation to predict the probability the loan will be repaid if originated. The predicted value of the dependent variable from the logit regression is the credit score, q. To be consistent with LaCour-Little (1996), the credit score is converted from the probability of repayment to the probability of default (by subtracting the credit score from 100 percent). This variable, NSCORE, is used as the dependent variable in the reverse regression. It is equivalent to the inverse of the qualifications index used in LaCour-Little (1996). The independent variables in the reverse regression are the

same as in LaCour-Little (1996); ACTION, which equals 1 if the loan was denied and 0 otherwise, and RACE, which equals 1 if the applicant is a minority and 0 otherwise. The logit regression model uses ACTION as the dependent variable, with NSCORE and RACE as the independent variables.

We use simulated data that contains 1000 observations on non-minority applicants and 500 observations on minority applicants. The average NSCORE for non-minority applicants is 26.4 percent with a standard deviation of 11.2 percent. The average NSCORE for minority applicants is 35.2 percent with a standard deviation of 13.4 percent. The first simulation contains no discrimination. All loans that have NSCORE's above 40 percent are denied and all other loans are approved. The rejection rate for non-minority applicants is 11.3 percent and the rejection rate for minority applicants is 36.2 percent.

The second simulation assumes discrimination exists such that non-minority applicants are approved with NSCORE's of 41 or below while minority applicants must have NSCORE's of 40 or below to be approved. The rejection rate for non-minority applicants in this simulation is 9.9 percent while the minority rejection rate remains 36.2 percent.

The third simulation assumes a wider range of discrimination exists, where non-minority applicants are approved with NSCORE's of 45 or below and minority applicants are approved with NSCORE's of 40 or below. For the third simulation, the rejection rate for non-minority applicants decreases to 5.0 percent, while the minority rejection rate remains at 36.2 percent.

If the two methodologies are reliable, they should not detect any discrimination in the first set of simulation data, but should detect discrimination in the second and third sets of simulation data. The findings should not reflect reverse discrimination for any of the three simulations.

The validity of default rate analysis is examined by comparing the average NSCORE of accepted minority applicants to the average NSCORE of accepted non-minority applicants for each simulation. Becker (1993) and others argue that, if discrimination is occurring, minorities should have lower default rates than non-minorities. Since the NSCORE variable measures the probability of default, according to Becker's (1993) theory accepted minority applicants should have lower average NSCORE's than accepted non-minority applicants if discrimination is occurring.

Since the required credit score is the same for all applicants in the first simulation, Becker's (1993) theory implies that the average NSCORE's for minorities and non-minorities should be the same for the first simulation. The second and third simulations result in lower marginal requirements for non-minorities, so Becker's (1993) theory indicates that non-minorities should have higher average NSCORE's than minorities for the second and third simulations.

RESULTS OF THE SIMULATION ANALYSIS

Default Analysis

The average NSCORE's for accepted minority and accepted non-minority applicants are shown in Table 3. The average NSCORE for accepted minorities remains 27.28 percent for all three simulations because the cutoff point does not change for minority applicants. The average NSCORE of accepted non-minority applicants is 24 percent when there is no discrimination (first simulation), 24.25 percent when there is slight discrimination (second simulation), and 25.21 percent when there is significant discrimination (third simulation).

Table 3: Average NSCORE for Accepted Applicants

	Simulation 1	Simulation 2	Simulation 3
Minority	27.28	27.28	27.28
Non-Minority	24.00	24.25	25.21

Since the NSCORE is the probability of default, it is apparent that default rates for accepted minority applicants can be higher that default rates for accepted non-minority applicants even if discrimination is occurring. This finding is consistent with the Ferguson and Peters (1995) model. If the distribution of credit quality is higher for non-minority applicants than for minority applicants, default analysis provides no useful information on whether or not discrimination is occurring.

The Performance of the Logit and Reverse Regression Models in Identifying Discriminatory Lending

The results of the reverse regression, calculations of a*, and the results of the logit regression are shown in Table 4. Using the reverse regression methodology, the RACE coefficient is positive and significant in all three simulations, and a* ranges in value from -.078 in the third simulation to -.156 in the first simulation. These findings are consistent with the results obtained by LaCour-Little (1996), which led him to conclude that there was evidence of reverse discrimination.

Table 4: Results of the Simulation Analysis

Reverse Regression Results

Simulation 1: No Discrimination

Variable	Parameter Estimate	T-Statistic	P-Value
Intercept	.2397	84.7	<.0001

ACTION	.2167	36.6	<.0001
RACE	.0339	6.8	<.0001
a* = -.0339/.2167 = -.156	Adjusted R^2=.528		
Simulation 2: Slight Discrimination			
Variable	Parameter Estimate	T-Statistic	P-Value
Intercept	.2425	84.9	<.0001
ACTION	.2187	35.5	<.0001
RACE	.0303	5.9	<.0001
a* = -.0303/.2187 = -.143	Adjusted R^2=.514		
Simulation 3: Significant Discrimination			
Variable	Parameter Estimate	T-Statistic	P-Value
Intercept	.2529	84.7	<.0001
ACTION	.2255	30.7	<.0001
RACE	.0175	3.1	.0019
a* = -.0175/.2255 = -.078			
Adjusted R^2=.451			
Logit Regression Results			
Simulation 1: No Discrimination -2logL = 1461 w/2 df p<.0001			
Variable	Parameter Estimate	Chi-Square	P-Value
Intercept	167.8	23.85	<.0001
NSCORE	-418.8	23.81	<.0001
RACE	-0.8	0.01	.9133
Simulation 2: Slight Discrimination -2logL = 1410 w/2 df p<.0001			
Variable	Parameter Estimate	Chi-Square	P-Value
Intercept	111.8	35.46	<.0001
NSCORE	-272.9	35.42	<.0001
RACE	- 2.5	11.36	.0007
Simulation 3: Significant Discrimination -2logL = 1257 w/2 df p<.0001			
Variable	Parameter Estimate	Chi-Square	P-Value
Intercept	107.9	31.16	<.0001
NSCORE	-239.9	31.11	<.0001
RACE	- 11.9	28.79	<.0001

The logit regression model performs well in detecting discrimination. The coefficient on RACE is insignificant in the first simulation, where there is no discrimination. The RACE coefficient is significant at the .001 level for the second simulation, where there is a slight degree of discrimination, and the RACE coefficient is significant at the .0001 level for the third simulation, where there is significant discrimination.

From the simulation analysis, it is apparent that the reverse regression methodology adapted from labor discrimination research is inappropriate for use in mortgage research. The logit regression methodology clearly outperforms the reverse regression methodology in detecting discrimination.

CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH

The findings indicate that we cannot focus on default analysis to answer the question of whether or not discrimination is occurring. The findings also indicate that the reverse regression methodology is not appropriate for mortgage discrimination research and that a well-specified logit regression model may be used to identify one aspect of discriminatory lending. It is important to recognize that the logit model can only identify discrimination that occurs after the loan application has been completed. Discrimination in pre-screening applicants, or differences in the assistance provided to applicants in completing the application process cannot be measured using the single-equation logit methodology.

One recommendation for future research is to study the distributional characteristics of credit quality and attempt to measure the differences in the distribution of credit quality between minority and non-minority applicants. This will be a daunting task, but if the distributional properties are better understood, we can devise a better method for detecting mortgage discrimination.

REFERENCES

Becker, Gary S. (1993). The Evidence Against Banks Doesn't Prove Bias. *Business Week*, April 13.

Blank, Emily C., P. Venkatachalam,, L. McNeil, and R.D. Green (2005). Racial Discrimination in Mortgage Lending in Washington, D.C.: A Mixed Methods Approach. *The Review of Black Political Economy,* 33(2), 1-30.

Brimelow, P., and L. Spencer (1993). The Hidden Clue. *Forbes*, January 4.

Browne, L.E., and G. M.B. Tootell (1995). Mortgage Lending in Boston – A Response to the Critics. *New England Economic Review*, Sept./Oct., 53-78.

Carr, J. H., and I.F. Megbolugbe (1993). The Federal Reserve Bank of Boston Study on Mortgage Lending Revisited. *Journal of Housing Research*, 4(2), 277-313.

England R. S. (1993). Washington's New Numbers Game. *Mortgage Banking*, September , 38-54.

Ferguson, M. F., and S. R. Peters (1995). What Constitutes Evidence of Discrimination in Lending? *Journal of Finance*, 50, 739-748.

Galster, G. C. (1993) The Facts of Lending Discrimination Cannot Be Argued Away by Examining Default Rates. *Housing Policy Debate*, 4(1), 141-146.

Horne, D.K. (1994). Evaluating the Role of Race in Mortgage Lending. *FDIC Banking Review*, 7(1), 1-15.

LaCour-Little, M. (1996). Application of Reverse Regression to Boston Federal Reserve Data Refutes Claims of Discrimination, *Journal of Real Estate Research*, 11, 1-12.

Liebowitz, S. (1993). A Study That Deserves No Credit. *Wall Street Journal,* September 1, A14.

Munnell, A.H., G.M.B. Tootell, L.E. Browne, and J.E. McEneaney (1996). Mortgage Lending In Boston: Interpreting HMDA Data," *American Economic Review*, 86, 25-53.

Rachlis, M.B., and A.M.J. Yezer (1993). Serious Flaws in Statistical Tests for Discrimination in Mortgage Markets. *Journal of Housing Research*, 4(2), 315-336.

Tootell, G.M.B. (1996). Redlining in Boston: Do Mortgage Lenders Discriminate Against Minority Neighborhoods? *Quarterly Journal of Economics*, November, 1049-1079.

Yezer, A., R. Phillips, and R. Trost (1994). Bias in Estimates of Discrimination And Default in Mortgage Lending: The Effects of Simultaneity and Self-Selection. *Journal of Real Estate Finance and Housing*, 9, 197-215.

USING COMMUNICATION THEORY TO ANALYZE CORPORATE REPORTING STRATEGIES IN THE BANKING INDUSTRY

Sheri Erickson, Minnesota State University Moorhead
Marsha Weber, Minnesota State University Moorhead
Joann Segovia, Winona State University

ABSTRACT

The purpose of this study is to investigate one specific industry, banking, and its management responses to material weaknesses in internal control within its Sarbanes-Oxley Section 404 reporting. Analysis of corporate financial disclosures using communication theory can provide useful information to stakeholders.

Benoit's (1995) image restoration typology assists in the determination of the communication strategies banks use to explain how such material weaknesses occur and how management intends to address those weaknesses. Because the banking industry has historically been highly regulated, one would expect few internal control problems or weaknesses, but results of our study indicate that this is not the case. Firms within the banking industry do report material weaknesses and these firms do not consistently indicate a corrective action strategy to remediate these weaknesses. Statistical results indicate that material weakness firms display higher market risk, greater asset growth, lower risk-based capital ratios, and riskier loan portfolios than firms in the banking industry. We also provide examples from corporate SEC and annual reports of banking companies to illustrate the use of various communication strategies based on Benoit's typology.

INTRODUCTION

The problems relating to business failures early in this decade exposed manipulations of financial reporting, distortions in economic performance in the accounting for and disclosure of transactions, and lapses in corporate governance. These problems resulted in Congress establishing requirements for corporate governance through its passage of the Sarbanes-Oxley Act (SOX), which requires firms to disclose material weaknesses in internal controls for financial reporting, directs management to disclose its assessment of those internal controls, and mandates that each company's independent auditor assess the management report and the company's systems of internal control. With the enactment of SOX, the U.S. Congress acknowledged major issues relating to the quality of earnings, transparency of financial reporting, and investor confidence in financial reporting and

directed the SEC to study a principles-based accounting system (United States Congress 2002, Sarbanes-Oxley Act Section 108). The major objective of SOX is to attempt to protect investors by improving the accuracy and the reliability of corporate disclosures that increase the transparency of reporting.

Recent failures in the banking industry and investigations into possible fraudulent activity in several large, failed financial service firms may indicate that the provisions of SOX have not gone far enough to prevent serious failures in corporate governance that can result in significant financial losses. The purpose of this study is to investigate one specific industry, banking, and its management responses to material weaknesses in internal control within its Section 404 reporting. Because the banking industry has historically been highly regulated, one would expect few internal control problems or weaknesses, but results of our study indicate that this is not the case. Firms within the banking industry do report material weaknesses and these firms do not consistently indicate a corrective action strategy to remediate these weaknesses.

This study uses Benoit's (1995) image restoration theory to identify the type of communication strategies banking companies employ to inform the public about material weaknesses found in internal controls. We also provide examples of Section 404 reports to illustrate the use of Benoit's theory to analyze management's responses. By studying these responses, we can determine whether management reacts to material weaknesses with corrective action strategies or whether other strategies are used. If management uses communication strategies other than corrective action, stakeholders may have concerns about management's acceptance of its responsibility for establishing, maintaining, and reporting on the effectiveness of internal controls and whether the company will address and remedy these internal control weaknesses.

Our results include an analysis of management responses within 65 banking companies that reported a total of 97 material weaknesses in their 2005 SEC filings (10-Q and 10-K). Benoit's (1995) image restoration typology assists in the determination of the type of communication strategies management uses to explain how such material weaknesses occur and what strategies management intends to use to address those weaknesses. As anticipated, the results of the analysis show that the majority of companies use a corrective action strategy when addressing material weaknesses. However, several firms use other communication strategies that include denial, evasion of responsibility, and reducing the offensiveness of the problem. When companies use communication strategies other than corrective action, management reporting is potentially not transparent and may indicate that management is not taking full responsibility for implementing and maintaining effective internal controls.

BACKGROUND

SOX requirements relating to financial reporting and internal control analysis emphasize management responsibility for preparing financial reports. This has generated academic interest, resulting in several different research streams regarding firm reporting and compliance under

Section 404, including the market reaction to internal control weakness disclosures (Hammersley et al. 2008; Beneish et al. 2006; De Franco et al. 2005) and the characteristics of firms that report material weaknesses (Doyle et al. 2007; Ashbaugh-Skaife et al. 2007; Ge and McVay 2005; Bryan and Lilien 2005). This study concentrates on how management communicates and addresses material weaknesses by analyzing firms' Section 404 disclosures within the banking industry.

We selected the banking industry for several reasons. First, more stakeholders exist in bank governance than in non-financial types of businesses due to banks' role in promoting the stability of the economy and the liquidity function. Therefore, loss of confidence in the banking system can cause serious economic problems and stakeholders should be concerned about the quality and transparency of financial reporting (Craig 2004). Second, Rezaee and Jain (2005) and Beasley et al. (2000) cite numerous studies that find that the financial services industries have a prevalence of fraudulent financial statements even though this is a highly regulated industry. In addition, studies by Loebbecke et al. (1989) and Akhigbe and Martin (2008) reinforce the importance of strong internal controls over top management.

Finally, we selected the banking industry because it has been subject to a higher degree of internal control regulation than nearly all other industries since the passage of the Federal Deposit Insurance Corporation Improvement Act (FDICIA) of 1991. Under FDICIA, all U.S. banks with total assets of $500 million or more are required to file an annual report with regulators in which management attests to the effectiveness of the bank's internal controls and an independent public accountant must attest to and report on management's assertions. Though banks have been subject to FDICIA requirements since the early 1990s, lapses in internal control exist. In 2003, the Basel Committee released a paper, *Sound Practices for the Management and Supervision of Operational Risk*, which outlines a set of principles that govern the management of operational risk at depository institutions. The Committee attributes many of the bank failures in recent years to operations risk in the form of executive fraud by manipulating internal controls, lax controls of information systems, or lack of financial competence of managers and boards of directors to monitor risk exposures. Operations risk may arise from inadequate or failed internal processes, people or systems, or from external events.

Internal control weaknesses are clearly a concern in the banking industry and SOX 404 requires that management addresses internal control weaknesses in reports to stakeholders. Image restoration communication strategies can provide insight into how these weaknesses are communicated by management and how they intend to correct those weaknesses.

Communication Theory and Image Restoration Strategies

Communication is a goal-directed activity that involves a purpose. One of the central goals of communication for the corporation is to maintain a positive image (Benoit, 1995). This is reinforced by the existence of public relations departments or public relations firms hired for the purpose of making or re-making company images. A reputation may be damaged intentionally or

unintentionally through word or deed. When this happens the communicator is faced with the problem of negative public image. Benoit creates his theory based on the assumption that, due to this potential negative image, the communicator is motivated to restore its image as one of the central goals of its communication to the public.

Business communication researchers have examined narrative portions of annual reports, including the CEO letter to shareholders and management's discussion and analysis and have begun to use communication theory and image restoration research to look at how managers and corporations communicate their financial and non-financial information to the public (Hildebrandt and Snyder, 1981; Thomas, 1997; Crombie and Samujh, 1999; Jameson, 2000; Rutherford, 2005; Deumes, 2008; and Lawrence and Geppert, 2008). Accounting and finance researchers are also beginning to use communication theory models in the study of corporate disclosure of material weaknesses in internal control in Section 404 reports (Erickson et al., 2009; 2008). Erickson et al. (2009) use Benoit's image restoration typology to determine what types of communication strategies computer firms use to disclose their material weaknesses in internal control and analyze what types of material weaknesses are associated with the use of non-corrective action strategies. Erickson et al. (2008) extend their analysis to examine whether the use of non-corrective action strategies is a red flag associated with a higher likelihood of merger, bankruptcy, or other type of serious financial difficulty within the computer industry.

Benoit's Image Restoration Typology

Benoit's (1995, 1997) typology is most often used by communication researchers to analyze strategic responses to legitimacy issues. Benoit's (1995) five categories of image restoration include denial, evasion of responsibility, reducing the offensive act, taking corrective action, and mortification. These five categories include fourteen unique communication (response) strategies as shown in Table 1.

Table 1. Benoit's Typology.

Categories	Strategy	Description/example
Denial	1. Simple denial	1 Refuting outright that the organization had any part in the event
	2. Shifting the blame	2 Asserting that someone else is responsible
Evasion of responsibility	3. Scapegoating	3 Blaming the event on the provocation of another
	4. Defeasibility	4 Not knowing what to do; lacking knowledge to act properly
	5. Accident	5 Claiming the event was "accidental"
	6. Good intentions	6 Claiming the company had good intentions

Table 1. Benoit's Typology.		
Categories	Strategy	Description/example
Reducing the offensive act	7. Image bolstering	7. Using puffery to build image
	8. Minimization	8 Stating the crisis is not bad
	9. Differentiation	9 Indicating that this crisis is different from more offensive crises
	10. Transcendence	10 Asserting good acts far outweigh the damage of this one crisis
	11. Reducing the credibility	11 Maintaining the accuser lacks credibility
	12. Compensation	12 Paying the victim; making restitution to set things to where they were before the event
Taking corrective action	13.Corrective action	13 Taking measures to prevent event from reoccurring
Mortification	14. Mortification	14. Admitting guilt and apologizing
Source: Benoit (1995)		

Denial can come in two forms. One is simple denial, or an outright refutation that the organization had any part in the event or was responsible in any way. The other type of denial is shifting the blame or asserting that someone (or something) else is responsible for the problem. Denial is the best strategy if the firm is truly blameless. If the firm uses a denial strategy and later is found to have blame in the event, its reputation can be irreparably damaged.

If denial is not an appropriate strategy, the organization can choose to evade responsibility by using one or more of four strategies. The first is scapegoating, which involves blaming the crisis on the provocation of another. Other evasion strategies include defeasibility, in which the organization did not know what to do or lacked the ability to act properly, claims the crisis was accidental or, that the organization had good intentions and therefore should be exonerated.

If a company cannot evade a responsibility that clearly exists, the company can reduce the offensiveness of the act by image bolstering (puffery), minimization (the event is not very bad), differentiation (this event was different from more offensive ones), transcendence (the good we do as an organization far outweighs the damage done by this one event), reducing the credibility of the accuser, and victim compensation. In the case of internal control weaknesses, victim compensation is not possible because no known loss has yet occurred. A lack of internal control is a warning that possible losses can occur in the future if such weaknesses are not detected and corrected in a timely manner.

Corrective action is Benoit's (1995, 1997) fourth strategy. An organization that uses this strategy tries to make amends for the wrong that was committed and takes measures to prevent the event from reoccurring. Corrective action is the most viable strategy in reporting internal control weaknesses because the firm addresses the source of the problem, explains how changes will eliminate future occurrences of the problem, and implements a remediation plan. When corrective action is used, management accepts its responsibility to maintain proper internal controls. The final

strategy proposed by Benoit is mortification, where the organization admits it was at fault and apologizes to the victims.

MANAGEMENT RESPONSE TO DISCLOSURE OF MATERIAL WEAKNESS IN INTERNAL CONTROL

As companies report their financial results, they strive to maintain a positive image and preserve legitimacy. A significant event or disclosure may result in a company trying to reduce the users' reaction to negative results. A company could also defend or restore its image to reduce the spotlight for misbehavior or wrongdoing. Thus, a company could "engage in recurrent patterns of communicative behavior designed to reduce, redress, or avoid damage to their reputation (or face or image) from perceived wrong doing" (Benoit, 1995, p.vii), a form of image restoration.

Benoit's (1995) image restoration typology assists in the determination of the communication strategies management uses. For example, SOX requires that company officers certify their responsibility for implementing adequate internal control policies and procedures. Because management is responsible for its assessment and evaluation of internal controls, we would expect management to take corrective actions when internal control weaknesses exist. However, management may use other strategies to disclose the weakness, to evaluate the weakness, or to address its responsibility to correct the weakness. Firms can use different or multiple communication strategies in explaining these weaknesses to their stakeholders.

Using Benoit's typology, one can gain insight into how a company reports internal control weaknesses under SOX 404 and assess management's responses to these weaknesses, which provides information on how the company intends to change and improve internal controls and whether it accepts responsibility for the weakness. The existence of internal control weaknesses can provide information about a potential pre-crisis situation. If the firm fails to correct the weakness, investor and creditor confidence in the firm's financial statements may be lost or the firm may even fail.

Management must first acknowledge why material weaknesses exist before they can correct them. By studying management responses, we gain insight into how serious management is about their internal control assessment and taking corrective actions to eliminate these weaknesses.

Previous research has addressed firms reporting material weaknesses in internal controls. One stream of research documents the characteristics of firms disclosing material weaknesses. Firm size is a determinant of good internal control (Ashbaugh-Skaife et al. 2007; Doyle et al. 2007; Ge and McVay 2005; and Bryan and Lilien 2005). Large firms are more likely to have more reporting processes in place and tend to have more employees and greater resources to spend on their internal control processes whereas small firms may lack sufficient resources to implement effective internal controls and may be more likely to use a non-corrective action strategy. Rapid-growth firms may outgrow their internal controls or they may dedicate a large portion of their resources to support growth rather than internal control processes (Doyle et al. 2007). Poorly performing (less profitable)

firms may not be able to invest in the proper internal control processes or they may be so concerned about improving their financial performance that they do not put sufficient resources and time into their internal controls (Ge and McVay 2005 and Ashbaugh-Skaife et al. 2007). The age of the firm may also be associated with the existence of material weaknesses as younger firms may not have the appropriate procedures in place to effectively manage their internal control processes leading to a higher likelihood of having material weaknesses and a lower likelihood of a corrective action response (Ge and McVay 2005; Doyle et al. 2007; and Ashbaugh-Skaife et al. 2007). Bryan and Lilien (2005) find that firms with higher market risk (beta) are more likely to have material weaknesses than firms with lower betas.

Another stream of research addresses types of material weaknesses firms disclose. Ge and McVay (2005) indicate that some of the most common deficiencies for banks are account reconciliation and period-end accounting policies. Deficiencies in accounting for loans (including accruals, loan loss reserves, provision for loan losses, and charge offs) may fall in both of these categories. We would expect that banks with riskier (lower quality) loan portfolios would be more likely to report material weaknesses than banks with higher quality loans. A bank's regulatory capital position may also affect its ability to respond appropriately to material weaknesses in internal control. Banks that do not meet regulatory guidelines may be concerned about improving their capital position and not devote the time and resources to improving their internal control systems.

The research has not evaluated how the firm communicates a material weaknesses and whether firms with different characteristics communicate the weaknesses differently. This discussion leads us to our first series of research questions:

RQ1: Do banking firms respond to material weaknesses with a corrective action strategy, recognize their responsibility, and take or plan to take action to correct the weaknesses?

RQ1b: How do material weakness firms in the banking industry compare to the banking industry as a whole in terms of size, profitability, age, growth, asset quality, market risk, and regulatory capital position?

Since the use of corrective action strategies implies that management acknowledges that they are responsible for the material weakness(es) and intends to make corrective changes, we are most interested in management's use of non-corrective action strategies. The use of these types of strategies could signal that management is less concerned with transparency in financial reporting and their strategic communication with the public.

Because the reporting requirements for SOX 404 are a result of Congressional action following corporate failures, the language used by firms to address material weaknesses in internal control is a relatively new type of corporate communication and firms were provided limited guidance on how to communicate this information. Firms may construct this communication in any

way they choose, so it is helpful to look at specific examples of how these messages are constructed. This discussion leads us to our second series of research questions:

RQ 2a: Do banking firms use non-corrective action strategies in SOX 404 reports to address material weakness deficiencies? If so, what are the most commonly used strategies?

RQ2b: If banking firms use non-corrective action strategies, what are specific examples of these strategies from our sample?

METHODOLOGY

This study uses a critical analysis method of studying communication strategies employed to repair tarnished images by carefully examining the language used by firms to communicate material weaknesses in internal controls and whether the company plans to correct the weakness in the future. An examination of the text of these communications provides insight into how companies use communication strategies to report these weaknesses.

Critical analysis of strategic communication has been used by many scholars, including Benoit (2006); Benoit and Czerwinski (1997); Benoit and Henson (2009); Blaney et al. (2002); Coombs (1995); Hearit (1995); and Seeger et al. (2003). A variety of texts have been evaluated using critical analysis, including speeches, advertising, newspaper articles, and public relations announcements.

We identified all firms in the banking industry (SIC Codes 6021, 6022, 6029, 6035, and 6036) that reported material weaknesses in internal control in their 2004 and 2005 10-K and 10-Q SEC filings from two sources: 1) Compliance Week and 2) EDGAR by searching the keywords "material weakness" and "internal control" and collected each firm's Section 404 report. These material weakness disclosures are made in Item 9A: Controls and Procedures of the firm's 10-K and in Management's Report on Internal Control over Financial Reporting and in Item 4: Controls and Procedures of the 10-Q.

Accelerated filers began reporting for years ending after November 2004 (2005 annual report). Thus, our data reflects the first time a firm reported an internal control weakness under Section 404.[1] If a firm reported no material weaknesses during 2005 but did indicate a correction for a material weakness stated in their voluntary reporting for the previous year, we collected the data for 2004 rather than 2005.

Two researchers, using critical analysis, read all the reports and independently classified the material weakness responses using Benoit's (1995) typology. Most firms' reports numbered the material weaknesses, followed by a corresponding number that contained the particular action the firm would take to correct the weakness. Any classification discrepancies between the two

researchers in classifying the responses according to Benoit's (1995) theory were discussed between the researchers and a consensus was reached as to the proper classification.

RESULTS

The resulting population includes 65 companies within the 773 firms in the banking industry that disclosed a total of 97 material weaknesses during the sample period. Some material weaknesses evoked more than one response, resulting in a total of 110 responses.

In response to research question 1a: Do banking firms respond to material weaknesses with a corrective action strategy, recognize their responsibility, and take or plan to take action to correct the weaknesses?, we find that management uses corrective action most frequently (97 times or 88% of the time). However, one of the most interesting observations is that some firms use image restoration communication strategies other than or in addition to the expected corrective action. The use of other strategies may provide additional insight into management's reporting that is not transparent and that management may not accept responsibility to take measures to prevent material weaknesses in the future. The use of specific non-corrective action strategies is explored further in the discussion for the second set of research questions.

In response to research question 1b, How do material weakness firms in the banking industry compare to the banking industry as a whole in terms of size, profitability, age, growth, asset quality, market risk, and regulatory capital position?, we used a univariate analysis to compare the median values of the material weakness firms and of the banking industry for firm size, profitability, growth, market risk, asset quality, and regulatory capital position. Table 2 shows the characteristics of material weakness firms (65 firms) versus the banking industry (773 firms) and the mean and median values for each variable. 2,3

Table 2: Characteristics of Material Weakness Firms versus Banking Industry 2005

	Material Weakness Firms		Banking Industry [1]		Wilcoxon Test Statistic (two-tailed) (p value)
	Mean	Median	Mean	Median	
VARIABLE					
Size	N=65		N= 773		
Total Assets (in millions)	$25,010.69	$1,775.82	$11,440.06	$848.33	0.0000***
Book Value (in millions)	$2,031.38	$147.88	$922.70	$77.52	0.0002***
Market Cap (in millions)	$3,971.85	$371.61	$1,766.19	141.46	0.0000***

Table 2: Characteristics of Material Weakness Firms versus Banking Industry 2005					
Employees	3.99	0.42	2.17	0.24	0.0001***
Profitability					
ROA (%)	0.98	1.01	0.93	0.94	0.0000***
NIM (%)	3.71	3.65	3.81	3.75	0.2670
Growth					
Asset growth (2000-2005) (%)	15.79	12.76	13.20	10.40	0.0000***
Market Risk					
Beta	0.70	0.78	0.30	0.27	0.0000***
Asset Quality					
Net charge-offs/Total Assets	-0.02	-0.01	-0.01	-0.01	0.0001***
RLL/Total Assets	1.19	0.81	0.80	0.80	0.0000***
PLL/Total Assets	0.26	0.18	0.20	0.10	0.0000***
Capital position					
Total Risk-Adjusted Capital (%)	13.59	12.21	14.58	13.02	0.0194**
Tier 1 Risk-Adjusted Capital (%)	11.12	10.64	11.89	11.00	0.1540

[1] SIC Codes 6021, 6022, 6029, 6035, and 6036.
***Indicates significance at the 0.01 level
** Indicates significance at the 0.05 level
All data were gathered from Compustat except Firm Age, which was gathered from Lexis Nexis. Total Assets, Book Value (equity), and Market Capitalization (price x shares outstanding) are dollar amounts shown in millions. ROA (%) is return on assets, measured by dividing net income before extraordinary items by total assets and is used as a measure of firm profitability. Another measure of profitability, NIM (%) is found by dividing net tax equivalent interest income by average interest earning assets. Employees indicates the number of employees per firm. Asset growth is calculated by finding the average annual growth rate in assets over a 5 year period. The asset quality variables are found by dividing net charge-offs, reserve for loan losses, and provision for loan losses by total assets, respectively. Total risk-adjusted capital represents the combined core and supplementary capital ratio and tier-1 risk adjusted capital is reported by equity capital plus minority interests less portion of perpetual preferred stock and goodwill as a percent of adjusted risk-weighted assets. Beta is a measure of market risk and is calculated over a 60- month time period. Firm age (in years) is number of years the firm has price data available on Lexis Nexis.
The Wilcoxon Statistic tests the statistical significance of differences in the medians of the material weakness firms and the banking industry. Medians were used instead of means because data for most firm characteristics are skewed.

Results of our study are consistent with those of Bryan and Lilien (2005), who find that firms with material weaknesses have higher levels of market risk (as measured by beta) than non-material weakness firms. Our study finds the median beta for the material weakness firms is 0.78, which is significantly higher than the median beta for the banking industry of 0.27. We also expect that

banks that report material weaknesses will have lower risk-adjusted capital ratios than firms in the banking industry. Results of our analysis indicate that this is the case; the Total Risk-Adjusted capital ratio is lower for the material weakness firms than for the industry as a whole. Also, as we anticipate, material weakness firms exhibit higher asset growth than the median for the industry. Finally, material weakness banks have higher medians for two of the asset quality ratios (reserve for loan losses/total assets and provision for loan losses/total assets) indicating that firms with material weaknesses may have riskier loan portfolios than the banking industry as a whole.

However, we also observed several differences from previous research. First, three size variables (book value, market capitalization, and total assets) for the material weakness firms indicate that material weakness firms are larger than the median for the banking industry. Another indicator of firm size, number of employees, indicates that the material weakness firms have more employees than the median for the industry. The return on assets, a measure of profitability, indicates that material weakness firms are more profitable than banking industry firms.

In response to research question 2a, Do firms use non-corrective action strategies in SOX 404 reports to address material weakness deficiencies? If so, what are the most commonly used strategies?, we find that 11 of the material weakness banking firms use other image restoration communication strategies, resulting in 13 non-corrective action responses. This may indicate that reporting lacks transparency and that management may not accept responsibility for these weaknesses or are not taking measures to prevent material weaknesses. Table 3 presents a summary of management strategies used to respond to material weaknesses.

Table 3: Image Restoration Strategies by Typology

Typology	Number of Times Used	Total for Category
Denial:		
Denial	1	
Shifting the Blame	0	**1**
Evasion of Responsibility		
Scapegoating	7	
Defeasibility	2	
Good Intentions	2	**11**
Reducing the Offensiveness		
Bolstering Image	1	
Minimization	0	
Differentiation	0	**1**
Taking Corrective Action		**97**
TOTAL		**110**

The most prevalent non-corrective action strategy used by firms in our sample is scapegoating, an evasion of responsibility strategy indicating that management does not want to take all of the responsibility for failure to implement effective internal controls. Other strategies used by firms in our sample to a lesser degree include defeasibility, a claim of having good intentions, shifting the blame to another party, bolstering the firm's image, and denial.

To address research question 2b, If firms use non-corrective action strategies, what are specific examples of non-corrective action communication strategies from our sample?, we selected excerpts from 10-K, 10-Q, and corporate annual reports from the 13 instances in which firms used non-corrective strategies to illustrate various communication strategies managers use to respond to material weaknesses in internal control over financial reporting based on Benoit's (1995) typology. We chose these examples to illustrate several non-corrective action strategies used by firms in our sample. Table 4 contains these excerpts and an analysis of the strategies used to respond to material weaknesses in internal control.

Table 4: Excerpts from SOX 404 Communications and Classification According to Benoit's Typology

Company	Excerpt Explanation	Benoit's Typology
Bay View Capital	We were unable to assess the effectiveness of internal controls of one of BVAC's third-party service organizations. The internal control weakness is blamed on the third-party vendor	Evasion of Responsibility - Scapegoating
Main Street Banks Inc	Material weaknesses occurred due to "the inability of the company to fill accounting, treasury, and other financial related positions as many organizations were seeking similar expertise." The material weakness was blamed on inability to compete with others seeking the same type of employees and executives in the company	Evasion of Responsibility - Scapegoating
Popular, Inc.	"…our U.S. mortgage and consumer lending subsidiary**,** with the intent to sell in the secondary market…were incorrectly presented as cash flows related with investing activities, instead of operating activities". The firm blamed the weakness on the failure of their subsidiary to properly classify cash flows	Evasion of Responsibility - Scapegoating
Mountain Bank Holding Company	…"we had not correctly recorded the benefit obligation for our executive retirement plan. A third party vendor was providing the calculation…and the Company's review control failed to identify errors". Although the firm did admit they failed to discover the error, they place the cause of the error on the third party vendor.	Evasion of Responsibility - Scapegoating
Midwest Bank Holdings	"The determination of whether available-for-sale securities are other than-than-temporary involves substantial judgment". Also, "FASB delayed the effective date…" The firm denies it recorded valuation declines in investment securities improperly by claiming that judgment is involved and that the effective date for the accounting rule change shields them from blame.	Denial – Simple Denial

Table 4: Excerpts from SOX 404 Communications and Classification According to Benoit's Typology

Company	Excerpt Explanation	Benoit's Typology
Juniata Valley Financial Corp.	Because the Corporation was not aware until the first quarter of 2005 that the Corporation is an accelerated filer, the Corporation was not able to complete its documentation and testing of internal controls in a timely manner. The company is claiming it did not know any better – that it lacked knowledge about SOX reporting requirements	Evasion of responsibility - Defeasibility
Mid Penn Bancorp	The number of finance personnel is a factor of the size of the company and management's attempt at *maintaining an efficient structure.* Management claims that insufficient controls were simply due to trying to run the business efficiently.	Evasion of Responsibility – Good Intentions
Mid Penn Bancorp	The independent firm hired by management (to assess internal controls) did identify several control deficiencies, but did not identify any material weaknesses. Management, however, is aggregating the deficiencies and is reporting them as a material weakness. Management is claiming here that they combined less serious deficiencies into a more severe material weakness which requires reporting. By doing this, management is trying to give the impression that they are being harder on themselves than they need to be.	Reducing the Offensiveness - Bolstering

Source: Company 10-K and 10-Q reports. Data available from authors upon request.

Although most banks used corrective action communication strategies to address material weaknesses, there were 13 instances in which firms used non-corrective strategies. Scapegoating, a form of evading responsibility, was the most commonly used non-corrective strategy. Bay View Capital used this strategy when it blamed the lack of internal control on a third party vendor. A problem with such a strategy is that the bank is ultimately responsible for work done by such third parties and therefore, scapegoating is not a viable strategy. Another example of scapegoating occurred when MainStreet Banks, Inc. attributed its material weaknesses to its inability to compete with other companies because there was a shortage of qualified financial and accounting personnel. Popular, Inc. placed responsibility on a subsidiary for failing to properly classify cash flows and Mountain Bank Holding Company also blamed a third party vendor for failing to provide the correct calculation for its executive retirement plan. A problem with using scapegoating as a means of evading responsibility is that when a company implies that the material weaknesses is in part due to the lack of diligence of another party, it indicates a lack of willingness to take full responsibility for its mistakes. This can lead stakeholders to question whether management is capable of overseeing the operations of the firm.

Mid Penn used another form of evasion of responsibility strategy, good intentions, when the firm claimed that the controls were inadequate because management was attempting to run the business efficiently. Its hope was that stakeholders would find this excuse admirable because all businesses should try to operate as efficiently as possible. Similarly, the evasion strategy of defeasibility, used by Juniata Valley Financial Corporation, is not a preferred strategy that management should use to avoid responsibility. When management attributed their inability to

correct internal weaknesses by claiming they lacked knowledge about SOX reporting requirements they indicated they were not up to the task of managing the company.

Mid Penn also attempted to reduce the offensiveness of its internal control weaknesses by bolstering its image. The bank claimed that the firm hired by management to assess internal controls listed several deficiencies, which are less serious and need not be reported in Section 404 reports, but failed to find any material weaknesses. Mid Penn combined these deficiencies into a more serious material weakness, thereby forcing the bank to report these findings in the report to stakeholders. By doing this, the bank is claiming that it have gone above and beyond their duties for reporting weaknesses.

Midwest Bank Holdings used the communication strategy of simple denial by asserting that financial statements were not misstated due to material weaknesses because the company had in fact followed current guidelines in reporting losses on investment securities. This strategy is effective if a company is truly blameless, but financial statements readers may be confused as to why this material weakness was identified by the external auditors if the firm was in fact correct in its assessment of the accounting rules.

An analysis of communication strategies used in SEC reports provides insight into management's reactions to internal control weaknesses and its use of corrective action to avoid a potential crisis. If firms use communication strategies other than corrective action such as denial; evasion of responsibility; or reducing the problem through image bolstering, minimizing the weakness, or stating that their situation is different from other companies, management may not accept responsibility in correcting internal control weaknesses. Users of the financial statements may have concerns about whether these internal control problems could lead to more serious financial difficulties and whether management is fairly reporting the company's economic reality. If management uses similar strategies in other financial reporting disclosures, the company's transparency in reporting may be questionable.

SUMMARY

Image management is essential to corporations and other organizations. If a firm is perceived by its stakeholders to be responsible for an event it performed, ordered, encouraged, facilitated, or permitted to occur, the firm's image will be tarnished and needs to be restored. Using Benoit's (1995) typology, this study provides evidence that when management reports internal weaknesses, banks are most likely to communicate their intended corrective action to eliminate these weaknesses in the future; however, our examination reveals that, in some instances, management uses strategies other than corrective action including denial, evasion of responsibility, and reducing the offensive act.

These non-corrective communication strategies could provide important insight to the users of financial statements. Strategies other than corrective action could signal to stakeholders that management may not be willing to take responsibility for correcting problems that created the

material weaknesses or that management is unwilling to establish and maintain disclosure controls over financial reporting as mandated by SOX. These strategies could reflect the overall control environment of the organization and provide a "red flag" of potential problems.

The statistical analysis in our study indicates that banks with material weaknesses have higher market risk and credit risk, greater asset growth, and lower risk-adjusted capital ratios than other firms in the banking industry. Results also suggest that material weakness banks are larger and more profitable than the banking industry as a whole.

FUTURE RESEARCH

Future research on the use of Benoit's (1995) typology to analyze corporate financial and non-financial disclosures could focus on several areas. First, do readers of management disclosures recognize the use of image restoration strategies? If so, do these tactics encourage a positive or negative reaction? Second, researchers could analyze whether certain combinations of image restoration strategies work more or less effectively than others. Finally, can Benoit's typology assist practitioners in designing messages during crises?

Additional research could also focus on other industries. While we focused on the banking industry, other industries could be studied to determine whether the same types of communication strategies were used to address material weaknesses. Future research could also focus on whether the communication strategies used to address material weaknesses had any relationship to whether those weaknesses were adequately addressed in future years.

Finally, in light of the current regulatory and financial issues affecting the banking and financial services industries, research can be conducted to determine what factors have caused failures and fraudulent activities, what effect the requirements of Sarbanes-Oxley have on the current situation, and how the problems can be remedied.

ENDNOTES

[1] Accelerated filers are companies with market capitalizations of at least $75 million, who have filed at least one annual report under Section 13(a) or 15(d) of the Exchange Act, and who are not eligible to file quarterly or annual reports on Forms 10-QSB or 10-KSB (SEC 2004).

[2] The distributions of these variables are highly skewed, so the Wilcoxon/Mann-Whitney Equality of Medians Test was used to compare the medians of the material weakness firms with those of the banking industry.

[3] A correlation analysis of the variables indicates high levels of correlation between variables used as proxies for the same characteristic (i.e. book value, market capitalization, and total assets are highly correlated and are all used as proxies for firm size). There are no other notable correlations between variables, including the correlation between ROA and NIM (both measures of profitability), which is 0.59.

REFERENCES

Ashbaugh-Skaife, H., D. Collins, & W. Kinney (2007). The discovery and reporting of internal control deficiencies prior to SOX-mandated audits. *Journal of Accounting and Economics,* 44 (1-2), 166-192.

Akhigbe, A. & A.D. Martin (2008). Influence of disclosure and governance on risk of US financial services firms following Sarbanes-Oxley. *Journal of Banking & Finance,* 32, 2124-2135.

Basel Committee on Banking Supervision (2003). *Sound Practices for the Management and Supervision of Operational Risk*. Basle, Switzerland: Basle Committee on Banking Supervision.

Beasley, M., J. Carcello, D. Hermanson, & P. Lapides (2000). Fraudulent financial reporting: consideration of industry traits and corporate governance mechanisms. *Accounting Horizons,* 14 (4), 441-454.

Beneish, M., M. Billings, & L. Hodder (2008). Internal control weaknesses and information uncertainty. *Accounting Review,* 83 (3), 665-703.

Benoit, W. L. (1995). *Accounts, excuses, and apologies: A theory of image restoration strategies.* Albany: State University of New York Press.

Benoit, W. L. (1997). Image repair discourse and crisis communication. *Public Relations Review, 23*(2), 177-186.

Benoit, W. L. (2006). President Bush's image repair effort on *Meet the Press:* the complexities of defeasibility. *Journal of Applied Communication Research, 34*(3), 285-306.

Benoit, W. L. & A. Czerwinski (1997). A critical analysis of USAir's image repair discourse. *Business Communication Quarterly, 60*(3), 38-57.

Benoit, W. L. & J.R. Henson (2009). President Bush's image repair discourse on Hurricane Katrina. *Public Relations Review, 35,* 40-46.

Blaney, J.R., W. L. Benoit, & L.M. Brazeal (2002). Blowout!: Firestone's image restoration campaign. *Public Relations Review, 28,* 379-392.

Bryan, S. and S. Lilien. (2005). Characteristics of firms with material weaknesses in internal control: an assessment of section 404 of Sarbanes-Oxley. Available at SSRN: http://ssrn.com/abstract=682363.

Coombs, W.T. (1995). Choosing the right words: the development of guidelines for the selection of the "appropriate" response strategies. *Management Communication Quarterly, 8,* 447-475.

Craig, V. (2004). The changing corporate governance environment: implications for the banking industry. *FDIC Banking Review,* 16 (4), 121-135.

Crombie, W. & H. Samujh (1999). Negative messages as strategic communication: A case study of a New Zealand company's annual executive letter. *Journal of Business Communication,* 36(3), 229-246.

De Franco, G., Y. Guan, & H. Lu (2005). The wealth change and redistribution effects of Sarbanes-Oxley internal control disclosures. Available for download at SSRN: http://ssrn.com/abstract=706701.

Deumes, R. (2008). Corporate risk reporting: A content analysis of narrative risk disclosures in prospectuses. *Journal of Business Communication,* 45 (2), 120-57.

Doyle, J., W. Ge, & S. McVay (2007). Determinants of weaknesses in internal control over financial reporting. *Journal of Accounting and Economics,* 44 (1-2), 166-192.

Erickson, S.E., M. Weber, J. Segovia, & D. Dudney (2009). Management Use of Image Restoration Strategies to Address SOX 404 Material Weakness, forthcoming in *Academy of Accounting and Financial Studies* Journal.

Erickson, S.E., M. Weber, J. Segovia, & D. Dudney. (2008, August). Can Section 404 material weakness communications predict mergers and SEC reporting problems? Poster session presented at the annual meeting of the American Accounting Association, Anaheim, CA.

Ge, W. & S. McVay (2005). The disclosure of material weaknesses in internal control after the Sarbanes-Oxley Act. *Accounting Horizons,* 19 (3), 137-158.

Hammersley, J., L. Myers, & C. Shakespeare (2008). Market reactions to the disclosure of internal control weaknesses and to the characteristics of those weaknesses under Section 302 of the Sarbanes-Oxley Act of 2002. *Review of Accounting Studies,* 13, 141-165.

Hearit, K. M. (1995). From "we didn't do it" to "it's not our fault": the use of apologia in public relations crises. In W. N. Elwood (Ed.), *Public relations inquiry as rhetorical criticism: Case studies of corporate discourse and social influence* (pp. 117-131). Praeger Series in Political Communication. Westport, CT: Praeger.

Hildebrandt, H. & R. Snyder (1981). The Polyanna hypothesis in business writing: initial results, suggestions for research. *Journal of Business Communication*, 18(1), 5-15.

Jameson, D. (2000). Telling the investment story: a narrative analysis of shareholder reports. *Journal of Business Communication,* 37(1), 7-38.

Lawrence, J. & J. Geppert (2008). Predicting firm reputation through content analysis of shareholders' letter. *Corporate Reputation Review*, 11 (4), 285-307.

Loebbecke, J., M. Eining, & J. Willingham (1989). Auditors' experience with material irregularities: frequency, nature, and detectability. *Auditing: A Journal of Practice & Theory*, (Fall), 1-28.

Rezaee, Z. & P.K. Jain (2005). Industry-wide effects of the Sarbanes-Oxley Act of 2002. *Journal of Forensic Accounting,* (6), 147-162.

Rutherford, B. (2005). Genre analysis of corporate annual report narratives. *Journal of Business Communication*, 42(4), 349-378.

Seeger, M. W., T.L. Sellnow, & R.R. Ulmer (2003). *Communication and Organizational Crisis.* Westport, CT: Praeger.

Thomas, J. (1997). Discourse in the marketplace: the making of meaning in annual reports. *Journal of Business Communication*, 34(1), 49-61.

United States Congress, House of Representatives, 2002, Sarbanes-Oxley Act of 2002. 107th Congress, 2d Session. Washington: GPO, 2002. Available for download at http://news.findlaw.com/hdocs/docs/gwbush/sarbaneSarbOxley072302.pdf.

THE INTERSTATE BANKING DEBATE: A HISTORICAL PERSPECTIVE

Jill M. Hendrickson, University of St. Thomas

ABSTRACT

This paper carefully documents the historical evolution of the small banker and large banker position on interstate banking in the United States. Using empirical methods, the validity of these positions is tested. The results indicate that the small bankers' position that branching leads to reduced competition was valid. At the same time, the large bankers' position that branching improves bank stability was also found to be empirically credible.

INTRODUCTION

Today the commercial banking landscape in the United States is rather diverse as it is comprised of extremely large bank holding companies with established branches throughout the county, regional and mid-size banks also with established branch systems, and small community bankers who primarily service local markets. This has not always been the case. Rather, the historical banking landscape is one of thousands of small, unit banks. Students of U.S. bank history recognize that part of the reason for the proliferation of these small banks was the legal inability of most bankers to open both interstate and intrastate branches. Indeed, branch restrictions emerged soon after the birth of banking, in the late eighteenth century. Contrary to these restrictions, many economists, regulators, and bankers recognized very early in U.S. banking history that branching could add stability to the enterprise. Ultimately, it took well over two hundred years for the United States to allow banks to branch freely. During that two hundred plus year process, there were many attempts to get regulators and legislators to change branching laws. Who or what was behind efforts to reform branching restrictions? Why did so many efforts fail? What case was made to preserve the branching ban and what case was made to eliminate it? Is there evidence that one side had a better case than the other?

While these questions, and their answers, are perhaps of interest to the economic historian, they should also be of interest to contemporary policymakers. In light of the 2007 - 2008 financial crisis and the government's role in its resolution, there is intense interest on how policy and regulation towards financial institutions, including commercial banks, will be altered in an attempt to avoid a similar crisis in the future. It is not known if these discussions will include interstate banking specifically but they will certainly include proposals to limit acceptable bank activity. In this way, the contemporary policy discussions are sure to return to discussions had hundreds of years

ago when regulation was enacted to keep banks from competing with one another and with other financial institutions. Another certainty is that bankers of all sorts and sizes will have an incentive to influence the outcome of the policy debates.

From this perspective, this research may be helpful to policymakers and economic historians alike because, to answer the questions above, it identifies and then empirically examines the commercial bankers' position on the branch banking issue throughout history. Though there are interested parties outside of banking, for example, non-financial businesses, the public, and regulators, to establish all positions is beyond the scope of this paper (see Kroszner and Strahan (1999) who consider the economic and political interest groups in relation to the deregulation of branch banking). Following this introduction, there are two main sections of this paper. First, this paper breaks the history of branch banking into four eras and attempts to establish evidence of the small and large banker position on the branching issue. Evidence of these positions comes from reading primary sources such as the *New York Times* and the *Wall Street Journal*, trade publications such as *Bankers' Magazine* and the *American Bankers Association Journal*, and government documents such as *Congressional Quarterly*. At the same time, this section of the paper also carefully documents market developments that certainly impacted the interest group positions in question. While this paper is not the first to address the history of branching in the United States, it is the first to document primary sources over the history of branching to establish interest group position. Almost all existing work takes as given the position of the small and large banker. This paper, in contrast, attempts to carefully document the positions from the beginning of the debate.

The second primary section of this paper attempts to empirically test the prominent interest group positions over the three eras for which data is available. Of interest is whether there is statistical evidence to support the soundness of one position or another and to determine if that evidence varied over the history of the debate. The empirical test finds evidence that both the small and large banker had statistically valid positions. More specifically, the analysis finds evidence that legitimizes the small bankers' fear of increased competition, while the large bankers' prediction that branching would enhance stability was also valid. To the author's knowledge, this is the first paper to examine these issues across the history of branching in the United States.

Since the beginning of commercial banking, the regulation of intrastate and interstate banking has attracted both critics and defenders. Interestingly, the arguments on both sides of the issue remained relatively consistent in the two hundred year debate. With time, some of the smaller points changed as the nation and banking changed. However, as this paper clearly illustrates, the earliest arguments for and against branch banking were still being voiced in the final debate in 1994.

Supporters of a branch system argue four fundamental points. First, they often argue that branch restrictions harm banks by limiting their ability to diversify (see, for example, Newfang (1901), Sprague (1903), Collins (1926), Wernette (1932), Jay (1933), White (1982), Wheelock (1992, 1995)). A unit bank is forced to extend loans and make investments in its immediate community that is frequently not too diverse. This means that an adverse development to the community could wipe out most of the bank's loans since its loan portfolio is the homogenous

reflection of the community. The assumption behind this argument both in theory and in practice is that banking markets are highly segmented. Bodenhorn (2003) provides several examples of banks during the antebellum era diversifying their portfolios by lending or investing beyond local markets. While this particular argument was leveled against branching restrictions up until the 1994 repeal, the homogeneity of community life was particularly acute during the antebellum and national bank eras. Branching restrictions also make a bank vulnerable on the liability side of its balance sheet. Like the assets, the liabilities of a unit bank cannot be too diversified which, in turn, magnifies the possibility of runs and gives banks a more narrow base from which to draw deposits.

Second, branch proponents argue that the branch restrictions, in our early history, kept banks from meeting the public's demand for depository services and constrained banks' ability to meet the financing needs of larger non-financial firms (see, for example, Newfang (1901), Sprague (1903), White (1982), Giedeman (2005), Preston (1924)). As the economy grew westward many communities were too small to support a national bank or even their own state bank but they needed banking services nonetheless. At the same time, growth in the non-financial sector resulted in larger and larger firms, most of whom required external financing. The contention is that commercial banks were too small, because of branch restrictions, to adequately provide large firm financing.

Third, those favoring a system of branches argue it would increase the competitive environment in banking and lead to lower priced loans and higher priced deposits to the benefit of the consumer. Branching would have the immediate effect of increasing competition as branch units compete with existing banks (Carlson and Mitchener (2006)). In theory, the increased competition would force weaker banks to exit leaving behind a stronger network of branches.

The fourth point often made by advocates of a branch system is that it may improve bank profitability and stability. Bank profit could increase through the more diverse balance sheets mentioned previously. Some scholars have also indicated that branching reduces administrative costs which could increase profitability (Southworth (1928)). Other scholars argue that branch systems afford banks access to each others deposits so fewer reserves are required which frees funds to be used for profitable purposes (Carlson (2004)). Improved profitability certainly enhances stability, but branch networks are seen to stabilize banking through another mechanism. If bank panics and runs are the result of asymmetric information problems so that depositors are unable to differentiate healthy from unhealthy banks, runs and panics ensue. A network of branches provides resources to banks caught up in the panic or run thereby reassuring depositors and restoring confidence (Bodenhorn (2003)).

On the other side of the issue are those determined to maintain a system of unit banks. Perhaps the most frequently stated reason for retaining the prohibition on interstate branching is the fear that it would lead to monopoly banking. As mentioned above, proponents of branching argue it increases the competitive environment to have branches competing against single office banks. However, opponents argue that if branching forces smaller banks to close their doors because they could not operate profitably, branching would have the effect of increasing industry concentration and reducing competition. A reduction in competition may mean a reduction in interest rates paid

on deposits and an increase in interest rates charged for loans (see, for example, Sprague (1903), Collins (1926), Wheelock (1992)). In essence, this position is concerned with the possibility of branching leading to a monopoly control of commercial banking.

Proponents of unit banking also argue that branching will hurt the small local borrower as the bigger, and more distant, bankers will perhaps be less willing to extend credit (Jay (1933)). Sprague (1903) identified early in the debate the asymmetric information problem that may arise with branching. Specifically, Sprague recognized the personal nature of many unit bank loans and the close relationship between the lender and borrower. This relationship provides the lender with more information about borrower risk characteristics. It is argued that some of this information may be lost, and important loans not made, under a branching system. Collins (1926) identifies this as the fundamental problem with a branch system.

THE EVOLUTION OF THE INTEREST GROUP POSITIONS

This paper chronologically analyzes the evolution of the branch banking debate in the United States. The analysis begins by attempting to identify the interest group positions for four periods in U.S. banking history: the antebellum era which considers the period from the inception of banking to the civil war; the national bank era which runs from 1864 to 1913; the Great Depression era which includes the years 1914 through 1935; and the modern era which picks up the analysis in the early 1960s through the passage of the Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994 that allows for interstate banking in the US effective June 1997. In addition to analyzing the interest group positions for these historical periods, market developments are also identified and analyzed. The empirical testing of the positions follows in the next section of the paper.

Antebellum Era

The state of commercial banking varied tremendously throughout the country in antebellum America. This variation was a reflection of the diverse degrees of economic development. Not surprisingly, branch banking regulation also varied from state to state. Chapman and Westerfield (1942) provide the most thorough consideration of branching during the antebellum era and find that most states actually allowed for branching, either intrastate or intracounty.

The antebellum period did see two cases of interstate branching as both the First and Second Banks of the United States were permitted to branch. That is, while state banks were prohibited from interstate branching, both federal banks established networks of branches across the states affording them all the advantages of branching. The First Bank of the United States (1791-1811) operated eight branches and the Second Bank of the United States (1816-1836) operated twenty eight branches (Boeckel (1937)). Thus, banks chartered by state authority were prohibited from interstate branching while the two banks with federal government charters enjoyed the privilege.

Not surprisingly, this discrepancy, particularly with the Second Bank, did not sit well with many and thus the branch bank controversy, on a national scale, was born in the United States.

The *Congressional Globe* documents congressional activity between 1833 and 1873. As such, it is the appropriate source for government discussion of branch banking during the antebellum era. There is no discussion directed specifically to branching. Rather, the vast majority of banking debate surrounded three issues: 1) the constitutionality and charter of the Second Bank of the United States; 2) currency activity of state banks; and 3) the payment of interest on deposits at state banks. The only specific reference found to branching was in 1836 when Senator Benton spoke against a proposed bill that would give authority to the Treasury to establish branches of state banks in states without banking (*Congressional Globe*, June 18, 1836, 500). Seven states – Arkansas, California, Florida, Iowa, Oregon, Texas, and Wisconsin – prohibited banking at some point in antebellum America (Hammond (1957)). Like many issues of the day, this branching idea was criticized as an intrusion of state rights. In the end, it appears Congress did not consider the branch issue during the antebellum era.

The *New York Times* historical database begins in 1851 and there is no evidence from this source that branch banking was a part of the national dialogue. Several other national sources also fail to identify branching as an issue during this era in banking history. Consequently, direct evidence of interest group positions in antebellum America on the branching issue has not been found.

Table 1 Evidence of the Small Banker Position

Argument	Evidence	Bank Era
1. Branching leads to monopoly banking.	1. Small bankers from Kansas proclaimed "we hereby affirm our unanswering allegiance to that view of the proposition which condemns it (branching) in all its forms as being unpatriotic, un-American, un-businesslike, and as tending to establish a monopoly of the great and honored business of banking in the hands of a few millionaires, to the exclusion of the men of the West, old and young, who have labored so faithfully and well to make our banking system what it is to-day, the best in the known world." (*New York Times*, May 15, 1902)	National Banking Era
	2. A group of small bankers from Missouri condemned branching as " . . it to be unwise, unsafe, unsound, and detrimental to the banking interests of the United States." (*New York Times*, May 15, 1902)	
	3. Declaration: "Resolved by the American Bankers' Association, That we view with alarm the establishment of branch banking in the United State and the attempt to permit and legalize branch banking; that we hereby express our disapproval of and opposition to branch-banking in any form by State or national banks in our nation."	Great Depression Era
	"Resolved that we regard branch banking or the establishment of additional offices by banks as detrimental to the best interests of the people of the United States. Branch banking is contrary to public policy, violates the basic principles of our	

Table 1 Evidence of the Small Banker Position		
Argument	Evidence	Bank Era
	Government and concentrates the credits of the nation and the power of money in the hands of a few." (*New York Times*, October 5, 1922, 8) 4. A small bank representative: "Mr. Frame, speaking first, said he would speak on "Monopoly vs. Democracy in Banking." Asserting that the 30,000 independent banks of the United States had done more to build the country than "all the cream-skimming monopolistic banks have done for other nations," he denied the contention of pro-branchers that branches served the people better or that failures and losses to depositors are lessened under branch banking." (*New York Times*, October 5, 1922, 8) 5. A representative of the American Bankers' Association warned that: "Branch banking will drive the small, independent community banker from America. . . and in their stead would spring up a system of hundreds of branch banks controlled by one or two leading national groups." (*New York Times*, November 7, 1923) 6. Excerpt from a small bank president: "The passage of the bill as it now stands, containing the branch bank feature, means the ultimate elimination of State and unit banks; it means the destruction of individual initiate and development, which is the thing that every American cherishes. . . The backbone of our country is the small, independent business and banking institution." (*Congressional Record*, 72nd Congress, 2nd Session, Vol. 76, Pt 2, 1468)	
2. Branching hurts the small borrower and local community.	1. "Although the American Bankers' Association, composed of a majority of small institutions, has taken a strong opposition to nationwide branch banking on the grounds that small communities are best served by local institutions…"(*New York Times*, October 28, 1927, 37). 2. "The fact that the local banker is the executive of final responsibility, that his greatest interest is in the community which the bank serves, that he is not restrained by the lifeless rule of bank bureaucracy, that he can, because of his intimate knowledge of clients, beam upon the spirit of petty local enterprise, all these are marshaled in support of the independent bank." (*Bankers' Magazine*, September 1930, 301) 3. A small community banker: "You'd have some guy in Atlanta deciding on whether to make a $50,000 mortgage in Mount Pleasant. He may decide that at that moment he'd rather put the money in Mexico for a better return." (Cocheo, April 1991). 4. "Interstate branching is bad public policy; it provides no benefit for community bankers; and it is little more that special interest legislation to allow our nation's biggest banks to consolidate their empires," according to an Independent Bankers Association of America statement. Consolidation, they claim, would cost communities jobs." (Dahl et al., 1995).	Great Depression Era Modern Era

Table 2 Evidence of the Large Banker Position		
Argument	Evidence	Bank Era
1. Restrictions limit balance sheet diversity.	1. Thomas Preston, the President of the Hamilton National Bank of Chattanooga, Tennessee, argued that branch banking "gives a diversification of loans and investments that is not possible in a unit bank, and in many instances a large bank with branches has been known to lend more in some small communities than the entire resources of that particular branch." (*New York Times*, August 1932, 15)	Great Depression Era
2. Restrictions limit banks' ability to meet client needs.	1. William A. Nash, President of the Corn Exchange National Bank, spoke before the New York Bankers' Association and he claimed that the greatest benefit of a branch system is the ability to move funds from one branch to another to meet demand for deposits and loans (*New York Times*, October 10, 1902, 16). 2. A President of a large bank spoke to Congress: "Banking, (too), has felt the pressure towards larger units and interconnections of units, the better to serve growing industries and communities. The quickened tempo of today's business has emphasized the interdependence of communities within the same natural industrial area, and had indicated the need for a more comprehensive and more closely knit banking service that has been available heretofore through the isolated unit banks." (Congress, House, Committee on Banking and Currency, Branch, Chain, and Group Banking, 71st Cong., 2nd Sess., April 1930, 1039)	National Banking Era Great Depression Era
3. Restrictions limit competition resulting in higher prices for consumers.	1. A banker representing the large bank interest commented: "In respect to branch banking our system is unique. The laws of every other important nation encourages branch banking, and the results of it have never tended to enslave the people, to build up dangerous monopolies, nor to increase the interest rate. The result, in fact, has been quite the reverse. Rates are kept uniform over a large territory, the tendency toward violent fluctuations is reduced, and the privileges and benefits of safe banking widely disseminated." (*New York Times*, December 14, 1907, 6) 2. A large bank representative observed: "Wherever a branch of our bank has been located where there was an existing local bank, it was permitted only after careful investigation by the banking department as to the need of another bank in that community. Whenever such a condition existed you will find in each case the then existing bank has today increased deposits equal to the total of our branch operating in that locality. Otherwise I contend a branch of our bank in that locality made a banking centre and not only stimulated business for itself but benefited its competitor." (*New York Times*, October 4, 1922, 7).	National Banking Era Great Depression Era

Table 2 Evidence of the Large Banker Position		
Argument	Evidence	Bank Era
	3. The CEO of Norwest Corp. observed: "It's (branch banking) about allowing banks to serve our customers wherever they are, wherever they want to be, and doing it faster, better and at a lower cost. No matter where they live, work, move or travel, consumers will be able to bank there, too." (Dahl et al. 1995).	Modern Era
4. Restrictions hurt profitability and bank stability.	1. The President of the First National Bank of Chicago, which was the second largest bank in the United States at the time, spoke before the Milwaukee Bankers' Club and indicated that a branching system would bring more profits because deposits could be used more efficiently (*New York Times*, May 5, 1902, 1).	National Banking Era
	2. A scholar observed: "The branch bankers, on the other hand, endeavored to prove that intercommunity branch banking had resulted in greater service to the people of the state and had contributed to the remarkable record of safety of banks in California during the past fourteen years. They favored permitting branches of national banks not only in the cities but on a basis of full equality with the state banks." (Preston 1924, 456)	Great Depression Era
	3. A Spokesperson for Northwest Bancoporation, the largest bank in the U.S. in 1930: "Communities will be insured greater financial stability and will be provided with broader and better services than has been possible by independent unit banks. Regardless of temporary depressions affecting one locality or industry, every bank will be able to render uniform and continuous service." (Neville 1930, 224).	
	4. Chairman and CEO of Banc One Corporation commented "[The interstate banking] legislation will lessen regional economic downturns, such as the one that hit New England several years ago. It is clear that in New England the downturn was made much worse because weakened banks were forced to shrink their loan portfolios as their capital levels fell off because of losses. Interstate banking, it is now recognized, would have enables banks to better withstand regional loan losses and to continue providing credit to job-creating businesses in New England." (*Wall Street Journal*, March 9, 1994, A12)	Modern Era

Tables 1 and 2 contain a sample of the interest group positions on the branch banking issue for the final three banking eras. Historical research indicates that the small banker was consistently opposed to branching on the grounds that it would create a monopolistic banking system at the expense of the local bank and community (see table 1). In the early history more emphasis was placed on the monopoly perspective and later more emphasis was placed on protecting communities and their bankers. The large banker was also consistent, through all three eras, advocating that branching was stabilizing and would benefit all bank clients (see table 2). In the earlier eras, the

large banker also emphasized the position that limits on branching made banks vulnerable to local economic downturns. Tables 1 and 2 paint a broader picture of the nature of the interest group positions throughout history. In what follows, more details of this evolution are provided for the final three banking eras.

National Banking Era

Financial historians refer to the national banking era as the period between the creation of nationally chartered banks in 1864 and the 1913 creation of the Federal Reserve. The opportunity to obtain national bank charters came with the passage of the 1863 National Currency Act and its subsequent revision as the National Bank Act of 1864. The National Bank Act was interpreted to prohibit the branching of national banks. That is, the act did not specifically prohibit branching, but the act required that all business take place at the location designated on the certificate of operation. The second Comptroller of the Currency, Freeman Clark, interpreted this to mean that national banks could not open branches (White (1983)).

Nationally chartered banks were not allowed to establish branches either within a state or between states. At the same time, many states prohibited branching for state banks and even in those states that afforded branching opportunities, few banks actually participated. Throughout the entire national banking era, there were fewer than 500 branches in the entire country (*Federal Reserve Bulletin*, various years).

The national banking era also witnessed five banking crises. While four of these were largely contained to New York and the surrounding area, the 1893 crisis was larger in terms of the number of bank failures (over 500) and in its geographic spread throughout the country. All of these experiences were certainly on the mind of those interested in the branching issue.

At the same time, U.S. banking witnessed a clear tendency of increased competition from foreign banks. In 1902, a writer for the *New York Times* described the growing trend for foreign banks to set up branches in cities other than the city of the parent bank. This happened because foreign banks were not subject to the branch restrictions of U.S. banks. He adds:

> *The fact appears to be that these changes are only another manifestation of the force impelling the financial business of the country as it has already impelled industrial business in the direction of larger organization and co-operation. Americans are as far behind the rest of the active business nations in their banking system as they are ahead of the rest of the world in industrial organization. (New York Times, May 19, 1902)*

The domestic crises and growth of foreign banks both added fuel to the branching debate during the national banking era. Yet, in the end, branching made little headway during this period, largely because of the political strength of the small banker.

Great Depression Era

Branch banking, as an important public issue, fell off the nation's collective radar following the national bank era. It was not until the early 1920s that branch banking came to the forefront of public attention. By the middle of the 1920s scholars were calling the branch issue the most important domestic bank issue facing the nation (Collins (1926)). Where did this interest come from? Much of it came from developments in the banking sector. For example, between 1920 and 1926, over two thousand commercial banks failed and most of these were small unit banks (Collins (1926)). One response to these failures was interest in a branch system over a unit system. Another development was the changing composition of commercial banks. Bradford (1930) indicates that beginning in about 1915, the growth in state banks began threatening the position of national banks. Moreover, Bradford (1930) points out that much of the state bank industry gains were the result of branch banking powers in certain states. Preston (1924) makes a similar claim to Bradford but is even more precise by arguing that four developments generated this renewed interest in branching.

First, the Comptroller of the Currency, D.R. Crissinger, observed that in some states many of the state banks were withdrawing from the national bank system so that they could branch. The Comptroller feared that the branching in some states represented an unfair advantage to those state bankers. In 1921, Mr. Crissinger asked Congress to grant nationally chartered banks the ability to branch within the city or county that the parent bank was located. When Congress failed to respond, the Comptroller took matters into his own hands. In early 1922 he let national bankers know that he would permit them the intra-city branch rights in those states that allowed branching. National banks responded by opening branches.

The second development was a market response to Mr. Crissinger's declaration. The First National Bank of St. Louis immediately established a branch within the city even though the state of Missouri prohibited branching. After a court battle, the bank was required to close its branch. Nonetheless, the controversy facilitated more national discussion on branching.

The third event was regulatory change by the Federal Reserve. The Federal Reserve Act allowed a state member bank the right to retain and perhaps even expand its branch system. However, regulators, much like the Comptroller, were becoming increasingly concerned that the state banks in branching states had a competitive advantage over nationally chartered banks. The 1924 revision to the act gave the Federal Reserve more authority in determining whether or not the state member bank could expand its branch system.

The fourth development that elevated branch banking to the national spotlight was a 1924 bill authored by Chairman of the House Committee on Banking and Currency, Mr. McFadden. The primary emphasis of this bill was the intrastate branching rights of state and national banks. This bill, in a revised form, would later become the 1927 McFadden Act whose details are outlined below.

Though not mentioned explicitly by Preston (1924), another development during this era that certainly contributed to the branch banking discussion was the tremendous growth in commercial

banks and, paradoxically, the increased industry concentration as the number of mergers increased throughout the 1920s. In 1920 there were over 22,000 state banks and over 8,000 national banks. At the end of the decade, there were 7,530 nationally chartered and 17,440 state chartered banks. Some of the declining numbers may be explained by the significant wave of mergers during this period. Between 1920 and 1930 there were 4,101 bank mergers in the United States (White, (1985)).

In addition to the growth and consolidation during the 1920s, there were three important regulatory developments during this decade. As mentioned above, in 1918 the federal government passed a law that allowed national banks to merge with one another. However, the act did not allow for the easy merging of national banks with other banking institutions. Indeed, if a national bank wanted to merge with a state bank it had to first convert to a state charter, merge, and then be a merged state bank. Since state banks had more intrastate flexibility in terms of branching (in some states) and fewer regulations, the number of national charters fell (in some states) as they converted to state charters.

The second regulatory development in the 1920s was the 1927 passage of the McFadden Act. Six years later, in 1933, Congress adapted further significant bank legislation as the banking sector was devastated by thousands of failures. Both of the regulatory developments are important to the branch banking history of the United States and are discussed next.

Alarmed by the number of conversions, in 1927 Congress passed the McFadden Act which allowed national banks to merge with state banks. The McFadden Act of 1927 also addressed issues of intrastate and interstate branching. The act allowed national banks to branch within the cities in which they were located, if state law permitted. However, the McFadden act effectively prohibited intrastate branching because a national bank could not open branches throughout the state as could state charted banks. In 1927, and for many years thereafter, states prohibited interstate branching for state chartered banks and the 1927 McFadden Act extended that prohibition to nationally chartered banks. Thus, the bank regulation of the 1920s on the one hand, allowed for consolidation while, on the other hand, confined banks geographically by prohibiting intrastate and interstate branching for national banks.

In the 1920s the branch issue was brought to the forefront of national bank policy by two market developments. One was the rapid increase in branch banks coupled with the conversion of nationally chartered banks to state chartered banks in order to take advantage of branch opportunities. Between 1918 and 1926, 206 nationally chartered banks were converted to state chartered and between 1901 and 1926 the number of branch units increased from 60 to 2,233 (*New York Times*, February 6, 1926, 14). The second market development was the alarming rate of failure among the smaller unit banks. In 1924, 613 small banks failed and the following year there were 464 failures (*New York Times*, February 6, 1926, 14). These market developments certainly gave the pro branching group a bit of an advantage in the policy discussion. However, the political power of the small banker was too much to move policy beyond intracity branching.

Small bankers continued to oppose extending branch rights. The evidence suggests that the primary position of the small banker was that it would be too much competition and that the small banker would be eliminated leaving behind a monopoly banking system (see table 1). Because of the market developments mentioned above, by October of 1922 the *New York Times* indicated that the biggest issue facing the American Bankers' Association was that of branch banking (October 3, 1922, 6). At their annual meeting, the members of the Association decisively voted against branching based generally on the fear that branching would lead to the elimination of the small banker and the monopolization of banking by a few.

At the 1922 American Bankers' Association meetings, many bankers representing large bank interests were present to try and convince the association that branching could actually help all sizes of commercial banks. While the large bankers argued that branching improves efficiency and allows banks to better meet the needs of the non-financial sector, primarily agricultural business (see table 2). Many of the small banks were already at their limit on agricultural loans and larger banks, with more capital, could extend more to each industry. Further, the large bankers were eager to establish branching rights and argued it would enhance the safety and stability of banking. In general the large banker was eager to establish branch networks.

There were two significant changes within commercial banking that certainly contributed to the widespread branching discussions and legislative developments. First was the rapid pace of conversion of national to state charters in part because of the branch privileges of many of the state banks. The second development in the banking sector was the relative stability of branching over unit systems. Calomiris (1990) indicates that between 1921 and 1929 only 37 branching banks failed. He also points out that those who witnessed banking in the 1920s were struck by the stability of branch systems and that many states responded by changing their branching laws in response to these observations. More specifically, between 1924 and 1939, the number of full or partial branching states increased from 18 to 36; a clear majority (Calomiris (1990)).

Another market development that cannot be ignored was the invention and use of the automobile. Collins (1926) discusses how traffic congestion made it increasingly difficult for customers to reach their downtown banks. Customers began demanding easier access to the banks outside of the busy downtown. Bankers wanted to meet this demand by establishing intra-city branches and, in many cases, intrastate branches. Further, improvements in telephone networks during this period enabled branch managers to communicate more effectively and at a reduced cost. In this way, branching became an increasingly attractive way for banks to expand.

Despite both the industry consolidation and expansion of national bank branching rights during the 1920s, the 1930s were a terrible disaster for U.S. commercial banking. Over 10,000 banks failed between 1921 and 1931. The majority of these took place after the passage of the McFadden Act and over 80 percent of all failures were small, unit banks (Jay, 1933). Given the banking crisis, it is not surprising that regulatory reform was called upon. The Banking Act of 1933 is a landmark piece of commercial bank legislation in terms of altering the operational parameters of banks. Among other things, the 1933 Act expanded branching rights for nationally chartered

banks from the city level in the McFadden Act to the state level. The Banking Act of 1933, to a certain extent, liberalized branch banking laws for nationally chartered banks by allowing them the same branching rights, within state, as state banks.

During the late 1920s and early 1930s, the vast majority of commercial banks were single-office unit banks. As was the case in earlier years, their numeric strength and dominate control over most banker associations put the unit bankers in a position to pressure regulators and legislators. The unit banker feared that if national banks were allowed to establish branch offices in their communities, they would be unable to compete with the larger banks (see table 1). The argument was that the branch system of larger banks would distance the banker from the needs of the local communities thereby leaving the public in a disadvantaged and insecure position.

Using their strength in numbers and their voice in influential industry institutions, such as the American Bankers' Association, the small commercial banker fought to keep the larger banks from entering most markets though out the country. They were motivated, in large part, by the fear that they would be unable to compete in a branch system.

Most large bankers were in favor of a branch system as a means of providing stability and prosperity to commercial banking. Representatives of large commercial banks argued that the smaller banks lacked necessary capital and were unable to properly diversify in order to be profitable (see table 2). Further, many large bankers argued that the process of intrastate and interstate branching would offer the diversification levels necessary for prosperity. Large bank advocates fought for liberalized branching rights based on the argument that the financial community needed to adapt to the changing conditions of the business community. During the late 1920s, the United States witnessed impressive growth in technology, transportation, and energy services. Additionally, industrial enterprises were changing in structure, due to competitive pressure, by becoming highly integrated and hence larger.

In the end, the coalition of small bankers was able to successfully fend off any further erosion in interstate banking. While the large banker enjoyed some small victories in this era with the intra-city banking in 1927 and then intrastate banking in 1933, interstate banking continued to be off limits.

Modern Era

The post depression years were rather quiet for commercial banking and while the interstate banking issue was kept alive it was primarily in academic circles. No serious attempts at changing legislative constraints were considered after the depression legislation until the 1990s. There were pockets of national interest over this time frame but interstate banking was not the primary issue in legislative and regulatory talks. More specifically, as market developments revealed weakness and fragility in the 1960s banking sector, the branch issue once again captured the interest of policymakers but nothing came of it. Major banking legislation was implemented in the 1980s but much of this was not primarily related to interstate banking. It was not until the early 1990s that the

focus was directly on interstate banking. The following provides some detail of the evolution of the issue in the modern era.

The branch discussion resurfaced in the early 1960s when James Saxon, Comptroller of the Currency, called for greater branching opportunities for nationally chartered banks. Saxon was concerned that the limits on branching would hinder macroeconomic growth. As population growth shifted to suburban areas the Comptroller was concerned that the 1933 legislation was often an obstacle for meeting the banking needs of the changing demographics. This was particularly true in states that either prohibited branching or limited branching.

As word of Saxon's proposal spread, branch advocates and opponents spoke in all too familiar terms. Small bankers resumed the argument of their predecessors as they were concerned that more branching by national banks would lead to an increase in bank concentration by eliminating many of the small banks. Large bankers favored expanding branch opportunities by arguing branching was more efficient, lowered costs, improved profits and consequently improved the stability of banking.

In the 1980s, regulators and legislators again began discussing branching but not directly. Rather, the concern was over the growth of "non-bank banks". In January of 1986, the Supreme Court limited the Federal Reserve's regulatory authority over non-bank banks. The Bank Holding Company Act of 1956 defined a bank as an institution offering both demand deposits and commercial loans. Non-bank banks could offer one, but not both, services. In this way, they did not fit the legal definition of a bank and so could cross state lines. According to financial experts at the time, the court ruling essentially opened the door to interstate banking. Thus while Congress was, in the 1980s, actively engaged in attempts to make important changes to commercial banking, the focus was not directly on branch banking laws outside the impact that non-bank banks would have on the interstate banking ban.

By the early 1990s, regulators and legislators could no longer ignore the branching issue because of important developments in the market that effectively forced the hands of these policymakers. Faced with increasing bank failures, a depleted deposit insurance fund, a rising number of non-bank banks, and the increasing number and size of regional banking compacts, Congress began looking at the interstate banking ban once again. In 1991 the issue was moved closer to real legal change than the efforts of the 1980s but still fell short.

Unlike the 1980s when branch banking was discussed primarily through indirect channels, in 1991 legislators and regulators pushed hard for interstate branching rights. However, the branching proposal was tied to a proposal to reduce insurance authority previously granted to commercial banks. In the end, this kept legislative leaders from lifting the branching ban. By 1994, market developments progressed to the point where they could no longer be ignored. Further, a bill passed by Congress required large banks to buy, at least for a time, a bank in another state before opening branches. This acquisition requirement appealed to the small bankers because it added value to their franchise. In many cases, this provision was enough to get the small banker to retreat somewhat in their opposition to the bill (*New York Times*, September 14, 1994, D1).

The debate in the modern era largely saw the small banker make many familiar arguments against interstate banking. For example, many small bankers continued to fight on the grounds that they would be driven out by big banks entering the market (*New York Times*, September 30, 1991, D5) (see also table 1). However, one development does set the modern era apart from earlier eras in that some small bankers (although certainly the minority) recognized publically that interstate banking may actually have benefits. Evidence of this change was found both in the 1991 debate and again in 1994. For the first time in the interstate banking dialogue, some small bankers actually spoke out in what may be described as lukewarm acceptance of this idea. A columnist for the *American Bankers Association Banking Journal* wrote:

> *Legislation easing restrictions on interstate banking will promote continued consolidation in the banking industry, but the impact on community banks will not necessarily be negative. The number of US banks grew steadily as the country expanded westward and has been shrinking just as steadily since then. Community banks will not necessarily disappear in this consolidation, but they will have to become more competitive. (Smith (September 1994, 17)).*

A small banker from Missouri was also optimistic as early as 1991 when he commented "When this state went to statewide bank holding companies, people were saying we might as well hang up 'For Sale' signs. That was 20 years ago and we're doing better now than ever." (Cocheo (April 1991)).

The fact that some in the small bank community were speaking in relatively positive terms about the possibility of interstate branching was a significant turning point. Since the chartering of the Second Bank of the United States in 1816, the small banker strongly and unwaveringly opposed interstate banking. It is likely that their experience with interstate branching through bank holding companies, which did not prove to destroy them, gave them the confidence to recognize that there was a place for both small and large institutions in commercial banking. Further, the large number of bank failures in the early nineties led many bankers, both big and small, to realize that a more diverse banking system could have withstood the downturn in both real estate and energy that led many banks into trouble (Bacon, 1994).

The large bankers once again pushed for the right to establish branches across the country. They argued that this would greatly reduce their costs by eliminating duplicative officers and functions *(New York Times*, February 4, 1994, D14). In 1991 Bank of America indicated that branching would allow the bank to save millions of dollars by eliminating duplication in labor, capital, and functions (Cocheo (April 1991)). In 1993, the CEO of BankAmerica Corp. estimated the costs savings more specifically stating that interstate branching could save the bank $50 million annually (Cocheo (January 1993)). In addition to cost savings, large bankers advocated that interstate branching would strengthen the banking system (see table 2).

Perhaps the most important factor behind the legislative lifting of the interstate banking ban was developments in the market that took place outside the halls of Congress. By the 1990s, commercial banks had successfully pressed into what was essentially interstate banking. Consider, for example, that at the end of 1991, all but two states permitted out-of-state companies to own banks within their borders (*CQ Weekly Online* (January 4, 1992)). By the middle of 1993, only Hawaii prohibited bank ownership by out of state bank holding companies (LaWare (1993)). Thus states had made it clear that they accepted interstate branching by creating these regional banking compacts throughout the country. Indeed, by the fall of 1994, there were 242 multibank holding companies in the United States (Kane (1996)).

Another market development in the modern era that allowed banks to effectively by-pass the interstate branching ban was the birth of electronic banking. Initially, this took the form of establishing a large network of automatic teller machines that allowed customers to engage in limited banking activity across state lines. Further, some banks were beginning to offer computerized banking which could also cross state lines. Thus, by the time the debate came to an end and intrastate branching was legally recognized, commercial bankers had essentially moved beyond the issue by practicing branching despite the national limitations.

EMPIRICAL ANALYSIS OF INTEREST GROUP POSITIONS

As the previous section of this paper indicates, the interstate branching debate was a long one and the banking environment evolved in significant ways throughout the life of the debate. One relative constant, however, was the position of the interest groups and the rationale behind their position. This section of the paper attempts to empirically test the validity of these positions.

Much has been written about the impact of bank regulation on bank performance (see, for example, Barth et al. (2004), Calomiris (2000), Levin et al. (1999), Kaufman (1996)), on how interest groups influence the regulation and policy process (see, for example, Becker (1983), Stigler (1971), McChesney (1997), Heclo (1978), Kroszner and Strahan (1999)), and more specifically on how branch banking impacts bank performance (see, for example, Amos (1992), Jayaratne and Strahan (1999), Carlson and Mitchener (2006), Ramirez (2003)). This research draws, to some degree, from all these types of existing literature to formulate the models presented here. At the same time, this research extends the body of literature in several ways. First, this paper considers how branch banking impacts bank performance across several historical eras. In contrast, most previous work utilizes either Great Depression data or data from the 1980s forward. This is probably due largely to data availability as it is difficult to create a comprehensive, pre-depression data sample. Second, this research carefully establishes the interest group positions and then empirically tests these over time. This research finds support for the small bank contention that branching may produce less competitive outcomes. At the same time, there is evidence to support the large bankers' position that branching enhances stability.

This analysis draws on previous empirical work that considers how branching impacts performance or market structures, something that both the small and large banker tried to use to influence branch policy. Amos (1992) was interested in learning more about the relationship between branching and bank closures in different states. His ordinary least squares (OLS) specification finds limited evidence to suggest that states which completely restrict branching and states with intrastate branching had more bank failures in the 1980s. Cebula (1994) extended the analysis of Amos by controlling for the financial condition of banks, by creating a single dummy branch variable, and by extending the analysis from the 1980s into the early 1990s. Cebula also finds that states with limited branching experience fewer bank failures which is consistent with comments in Amos (1992). Loucks (1994) takes the Amos relationships but estimates them with the Tobit estimator rather than OLS because some of the dependent variables in Amos are truncated at zero. Louck finds a smaller percentage of bank closures in states with less branch restrictions. Ramirez (2003) takes a more micro approach to the analysis of branching and bank failures by comparing the experience of a branching state with a non-branching state in the 1920s. He finds evidence to indicate more bank failures associated with the unit banking environment. Further, Ramirez suggests that this is because branching reduces costs allowing for larger banks who enjoy more diversified balance sheets.

Other literature controls for the economic diversity of the area in which banks operate. Throughout history, the large banker consistently made the case that branching enhances the diversity of a bank's portfolio and, in doing so, makes the banking system more stable and profitable. Shiers (2002) creates a measure of economic diversity and empirically tests whether this diversity impacts bank risk and profits. For the period 1966 to 1981, Shiers finds that economic diversity reduces bank risk and he also finds similar findings for the time period 1982-1996. More recently, Carlson and Mitchener (2006) test how the growth in branching during the Great Depression influenced bank competition and how the competitive changes affect bank failures. Their hypothesis is that as branching expands, competition increases which leads to exit by banks less able to compete so the banking sector becomes more stable. The empirical findings of Carlson and Mitchener (2006) suggest that, for nationally chartered banks, the competition hypothesis contributed more to stability than did an increase in portfolio diversification. This suggests that the arguments for diversification were perhaps not as valid during the Great Depression era as were the fears of the small banker who anticipated they would be unable to compete with the bigger banks.

This study uses annual aggregate state data for an eleven year span within each banking era. For each banking era, the data is pooled across the states and all bank data includes both state and nationally chartered banks. Ideally this empirical investigation would break the commercial bank data (during the first two eras) into two groups: one for nationally chartered banks and another for state chartered banks. This would be extremely helpful since in these earlier years different chartered banks operated in different branching environments. For example, during the national banking era, some state banks were allowed to branch within state while national banks were not.

However, neither failure nor merger data at the state level for *both* charter types has been found for each year of the sample periods.

In terms of failure data, the *Annual Report* of the Comptroller of the Currency records the number of national bank failures by state. The number of state bank failures in each state, in contrast, is not available for each year of the national banking era. There are several sources in addition to the publications of the offices of the Comptroller that provide historical bank failure data. These include Goldenweiser (1933), *Banking and Monetary Statistics*, *All-Bank Statistics*, and historical issues of the *Federal Reserve Bulletin*. In the end, however, there is not currently available bank failure data by state and charter type for both the national banking era and Great Depression era.

In terms of bank merger data, like the failure data, there exists several data sources including Chapman (1934) and the *Federal Reserve Bulletin* (1937). While some of this historical merger data is at the state level, it does not break down into both charter types for both the earlier bank eras. Because of failure and merger data limitations, this paper uses state level banking data without distinguishing between charter types (see Appendix A for all data sources and definitions).

The national banking era sample covers 1900 to 1910, the Great Depression era sample runs from 1924 to 1934, and the modern era sample is from 1984 to 1994. These dates were chosen, in part, due to data constraints but also because of the timing of events such as regulatory changes or banking crises. For example, both the national banking era and Great Depression era experienced financial crises. The national banking era sample years were chosen because there was only one crisis during this time frame and no significant changes in bank regulation. It is nearly impossible to avoid bank crisis or regulatory change in the Great Depression era, unless the analysis ended in or before the 1927 passage of the McFadden Act. In this analysis, the 1933 crisis is controlled for using a dummy specification for the crisis year and the branching regulatory environment is also controlled for. In the modern era, the sample begins after significant bank regulatory change and ends with the passage of the Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994.

For each of the banking eras with identified interest group positions, three interest group positions are specified. First, the small banker unfailingly argued that an extension of branching rights would lead to increased market concentration and monopoly power. The first model attempts to capture the impact of branching on competition. Second, the large banker insisted that allowing for more liberalized branching would make the banking system more stable. Thus, model two captures the impact of branching on bank failures. Finally, the large banker consistently argued that branching would allow them to diversify their portfolios and, consequently, be able to withstand macroeconomic downturns. The failure model contains an explanatory variable meant to capture diversity. Consequently, inferences will be made regarding the relationship between the ability of banks to diversify and its impact on bank failures.

In constructing the models, it is necessary to control for the financial health of the banks (in the failure model), the market structure the banks are operating in, the regulatory environment

influencing banker behavior, and the economic health of the state. Model specifications and variable definitions follow.

Competition Model

The small banker spoke loudly and often throughout the branching debate and their message was consistent: the extension of bank branching will lead to less competition and even monopoly banking. To test the validity of this point of view, a competition model is constructed in which the dependent variable captures the extent to which each state banking market was competitive (see equation 1 where i represents the state and t represents the year):

$$(1)\quad COMPETITION_{it} = \beta_0 + \beta_1 BRANCHING(-1)_{it} + \beta_2 NUMBER(-1)_{it} + \beta_3 YEAR_t + \beta_4 CRISIS_t + \beta_5 STATEECONOMY(-1)_{it} + error$$

Following Carlson and Mitchener (2006) COMPETITION is defined as the number of bank mergers. Ideally, a measure of competition could be created that more directly measures competition. For example, a Herfindahl-Hirschman Index would be ideal. However, data constraints prohibit this. In terms of the independent variables, the model tests two specifications to control for the regulatory environment within each state (labeled BRANCHING(-1) in equation 1). One specification, BRANCH, utilizes the one period lag of the number of banks operating branches to control for the level of branching in each state. During the national banking era, if two national banks wanted to merge, one bank was liquidated while the other purchased its assets and liabilities. State banks could often merge, but could not merge with a national bank unless the liquidation process described between two national banks was employed. Alternatively, a state bank and national bank could merge by taking out a state charter. Because of these difficulties, there was not a significant amount of merger activity in this early banking era. In terms of the impact of branching, state banks did use merging as a means of establishing branches so we may expect a positive relationship between branching and merging. During the Great Depression era, national banks were allowed to merge and consolidate without liquidation provided the Comptroller approved. The hurdles to state and national bank merging were in place until the passage of the McFadden Act of 1927 which allowed national banks and state banks to merge under the same rules. As in the earlier era, many banks used mergers to establish branch networks. Consequently, it is expected that this measure of branch activity, BRANCH, to positively impact the degree of concentration in the banking market.

For the alternative branch specification, Shiers (2002) suggests UNIT, the lag of the ratio of the total number of unit banks to the total number of banks, as a measure of the extent of non-branching in each state. This specification captures the opposite of BRANCH and so is expected to have a negative impact on this measure of competition. These two alternative specifications, BRANCH and UNIT, used to control for branching are utilized in the all three eras. However, in

the modern era, more specific data indicating the date branch laws changed in each state are available (Stiroh and Strahan, 2003). Consequently, in the modern era, an additional regression is analyzed using controls for the legal changes in branch laws (results in table 6).

The remaining independent variables control for the state economy, the market structure, and time-specific effects. %BUSFAILURE and GSP control for the performance of the state economy (labeled STATEECONOMY(-1) in equation 1). During the two earliest eras, the state economy is proxied as the lag of the percent of business failures in each state. During the modern era, this is measured as the lag of the gross state product. Certainly, the healthier the state economy, the more likely the banking sector is also healthy. A strong economy may signal more opportunities for mergers as banks expand and have the financial ability to purchase other banks. Another possible scenario is one in which there will be more mergers when the state economy is struggling as banks find themselves in weaker positions. From this perspective, bank mergers may be a way for weaker banks to avoid failure. Consequently, the impact of the state control on merger activity is uncertain. To control for the market structure that the banks are operating within, this paper relies on the lag of the number of banks, NUMBER, operating in each state. An increase in the number of banks would lead one to expect fewer bank mergers since merging reduces the number of bank entities. Indeed, Berger, Kashyap, and Scalise (1995) find an increased reduction in the number of banks following the introduction of branch opportunities. Since more branching tends to come, in part, from merging, NUMBER is expected to negatively impact COMPETITION.

YEAR controls for time-specific effects. Given that, with the passage of time, merging became easier in the Great Depression and modern era, it is expected that YEAR will positively impact the number of mergers, particularly in the later two eras. CRISIS controls for any financial crises during the sample period.

Failure Model

The large banker long insisted that more opportunities for branching would improve competition and leave the more efficient and stronger performers standing. From this perspective, branching makes banking more stable, at least after the weaker banks have exited. To test the impact of branching on stability, this paper relies on a failure model since there should be fewer failures, after a period, with opportunities to branch:

$$(2)\quad FAILURE_{it} = \beta_0 + \beta_1 BRANCHING(-1)_{it} + \beta_2 NUMBER(-1)_{it} + \beta_3 YEAR_t + \beta_4 CRISIS_t + \beta_5 STATEECONOMY(-1)_{it} + \beta_6 CAPASSET(-1)_{it} + \beta_7 LOANASSET(-1)_{it} + \beta_8 DEPASSET(-1)_{it} + \beta_9 DIVERSITY(-1)_{it} + error$$

The dependent variable, FAILURE, is defined as the number of bank failures in each state. BRANCHING, NUMBER, STATEECONOMY, CRISIS, and YEAR are defined as in the COMPETITION model.

To control for the financial health of the banks, three ratios are utilized. The lag of the capital to asset ratio, CAPASSET, indicates the extent to which a bank is able to withstand unexpected adverse conditions (Ramirez (2003), Cebula, (1994)). A lower capital to asset ratio indicates more risk and consequently more failures. LOANASSET, the lag of the ratio of loans to assets, is an indicator of bank risk since loans are typically the most risky asset on a bank balance sheet. As suggested in Ramirez (2003), the higher the ratio of loans to assets, the greater the likelihood of failure. Deposits are usually the cheapest source of funds for a bank and, according to Ramirez (2003), banks with large deposit bases are seen to be financially stronger given the cost advantage from those deposits. Consequently, it is expected that the larger the lag of the ratio of deposits to assets, DEPASSET, the fewer failures. Finally, DIVERSITY is defined as the lag of the ratio of the total number of branches to total number of banks in each state. Carlson and Mitchener (2006) use a similar measure and point out that more opportunities to branch should improve balance sheet diversification and therefore reduce failures.

Given that the data in this research contains both cross-sectional (state) observations and time-series (year) observations, the estimation procedure must be appropriate to this panel data. There are several types of models for panel data. These include the constant coefficient models, fixed effects models, and random effects models. For this research, significant state differences are expected because of different branch laws so the constant coefficient model is not appropriate. To determine whether the fixed or random effects model should be used, it is common to use the Hausman specification test (Wooldridge (2002)). Doing so indicates that testing the impact of branching on competition and bank failures should be done utilizing the fixed effects estimation strategy. The correlation matrix detects no significant multicollinearity in the specifications. Further, each specification is tested for the presence of heteroskedasticity using the Breusch-Pagan test. However, the DW statistic indicates the presence of autocorrelation. This is corrected for utilizing an AR estimator.

Table 3: Estimation Results for COMPETITION in National Banking Era

Variable Name	Number of Banks Operating Branches		Ratio of Unit Banks to Total Banks	
	Regression 1	Regression 2	Regression 1	Regression 2
UNIT(-1)			-4.11*** (1.72)	-5.93** (1.97)
BRANCH (-1)	0.074*** (1.76)	0.154*** (1.81)		
NUMBER(-1)	-0.004** (2.06)	-0.001** (1.96)	-0.001** (1.94)	-0.003** (2.02)
YEAR	0.111** (2.08)	0.202*** (1.87)	0.128* (2.73)	0.126* (2.62)

Table 3: Estimation Results for COMPETITION in National Banking Era				
CRISIS	-0.578*** (1.68)	-0.869* (2.81)	-0.529** (2.08)	-0.64 (1.57)
%BUSFAILURE(-1)		-0.416 (1.26)		-0.475* (2.51)
Observations				
Cross-section	25	25	25	25
Time-series	225	225	225	225
R^2	.576	.647	.614	.643
F-statistic	6.48	6.719	6.219	6.289
DW statistic	2.067	2.035	2.081	2.085

Dependent Variable: number of mergers. Independent variables: "BRANCH" is the lag of the number of banks operating branches in each state for each year. "NUMBER" is the lag of the total number of banks in each state. "YEAR" controls for time-specific effects. "CRISIS" is 0 for all years without a financial crisis, 1 for 1907. "%BUSFAILURE" is the lag of the percent of commercial business failures in each state. "UNIT" is the lag of the ratio of total number of unit banks to total banks in each state. The first two columns of results are for the legal branch status control and the third and fourth columns are for the extent of branching control specification. Regression 1 contains a control for the regulatory environment and market structure, Regression 2 contains controls from (1) and adds a control for the state economy. t-value in parenthesis. *significant at 1% level. **significant at 5% level.***significant at 10% level. Note that data availability was limited for this era. Consequently, only 25 states are included in this analysis: Connecticut, Florida, Georgia, Idaho, Illinois, Indiana, Kansas, Louisiana, Michigan, Minnesota, Missouri, Nebraska, North Carolina, Ohio, Oregon, Pennsylvania, Rhode Island, South Carolina, South Dakota, Texas, Utah, West Virginia, Wisconsin, Wyoming. For data sources, see Appendix A.

.

Table 4: Estimation Results for COMPETITION in Great Depression Era				
Variable Name	Number of Banks Operating Branches		Ratio of Unit Banks to Total Banks	
	Regression 1	Regression 2	Regression 1	Regression 2
UNIT(-1)			-1.732*** (1.86)	-1.12*** (1.81)
BRANCH (-1)	0.097*** (1.69)	0.101*** (1.76)		
NUMBER(-1)	-0.001** (1.84)	-0.001** (2.01)	-0.003*** (1.81)	-0.002*** (1.84)
YEAR	-0.208** (1.89)	-0.216** (1.97)	-0.206*** (1.76)	-0.215*** (1.83)
CRISIS	0.714 (1.30)	0.688 (1.25)	0.67 (1.24)	0.652 (1.19)
%BUSFAILURE(-1)		-0.829 (1.07)		-0.741 (0.958)

Table 4: Estimation Results for COMPETITION in Great Depression Era				
Variable Name	Number of Banks Operating Branches		Ratio of Unit Banks to Total Banks	
	Regression 1	Regression 2	Regression 1	Regression 2
Observations				
Cross-section	47	47	47	47
Time-series	423	423	423	423
R^2	.664	.701	.686	.698
F-statistic	10.017	9.851	9.893	9.926
DW Statistic	2.107	2.107	2.102	2.102

Dependent Variable: number of mergers. Independent variables: "BRANCH" is the lag of the number of banks operating branches. "NUMBER" is the lag of the total number of banks in each state. "YEAR" controls for time-specific effects. "CRISIS" is 0 for all years without a financial crisis, 1 for 1933. "%BUSFAILURE" is the lag of the percent of commercial business failures in each state. "UNIT" is the lag of the ratio of total number of unit banks to total banks in each state. The first two columns of results are for the legal branch status control and the third and fourth columns are for the extent of branching control specification. Regression 1 contains a control for the regulatory environment and market structure, Regression 2 contains controls from (1) and adds a control for the state economy. t-value in parenthesis. *significant at 1% level. **significant at 5% level.***significant at 10% level. For data sources, see Appendix A.

RESULTS

Tables 3 through 6 contain the regression results for the fixed effects COMPETITION model for the national banking era, Great Depression era, and modern era respectively. These findings are consistent with existing literature that finds branching leads to more bank mergers. Specifically, Carlson and Mitchener (2006) find that states with branch systems witnessed more mergers during the Great Depression and White (1985) analyzed bank mergers in the 1920s and suggests that there is a positive relationship between branching opportunities and merger activity. The results in table 4 are consistent with this existing Great Depression era research and the results in table 3 indicate that branching in the pre-depression era also increased mergers, and consequently, consolidation in a state's banking system. The estimation results in table 5 are consistent: even in the modern era, the estimated coefficients for both branch specifications support the hypothesis that more branching leads to more banks mergers. These results validate the small bankers' position which was to consistently fight attempts to expand branching on the grounds that it would lead to highly concentrated markets.

The results also suggest that the performance of each state economy in the two early eras has a different influence on COMPETITION than the modern era. More specifically, in both the national banking era and Great Depression era the estimated coefficients on %BUSFAILURE indicate a reduction in bank mergers, i.e. more competition, during more unstable economic times. In contrast, the results in the modern era suggest that slower economic growth corresponds to more mergers, i.e., less competition. Yet, the results are only statistically significant in the modern era

and in one regression in the national banking era. Some of these differences may reflect different measures (the lag of the percent of commercial failures in the two early eras and the lag of gross state product in the modern era). Since gross state product is a more comprehensive measure of a state's economic health, the results of the modern era may be more indicative of the relationship between the state economy and level of bank merger activity.

These results suggest that the structure of the market also matters in terms of bank merger activity. In each regression, the control for market structure, NUMBER, is of the expected sign and statistically significant. This is consistent with existing research that finds merger activity to be related to the number of banks in the market.

Finally, the control for time-specific effects is statistically significant across the three eras. Further, the estimated coefficients indicate a positive relationship between time and merger activity, as expected. The control for bank crisis in the time series is statistically significant in the national banking era suggesting the crisis impacted merger activity but the crisis control was not significant in the Great Depression era.

Table 5: Estimation Results for COMPETITION in Modern Era

Variable Name	Number of Banks Operating Branches		Ratio of Unit Banks to Total Banks	
	Regression 1	Regression 2	Regression 1	Regression 2
UNIT(-1)			-1.144 (1.21)	-3.88 (1.41)
BRANCH(-1)	2.451** (2.05)	2.948** (2.11)		
NUMBER(-1)	-0.0002*** (1.86)	-0.0001** (2.16)	-0.0001** (2.04))	-0.001** (2.15)
YEAR	0.385*** (1.98)	0.825** (2.06)	0.199 (0.94)	0.792** (2.03)
GSP(-1)		-0.0002* (2.94)		-0.0001* (3.00)
Observations Cross-section Time-series	 50 450	 50 450	 50 450	 50 450
R^2 F-statistic DW Statistic	.604 10.428 2.045	.624 12.304 2.921	.593 10.894 2.059	.602 11.081 2.082

Table 5: Estimation Results for COMPETITION in Modern Era

Dependent Variable: number of mergers. Independent variables: "BRANCH" is the lag of the ratio of the number of banks operating branches to total banks in each state. "NUMBER" is the lag of the total number of banks in each state. "YEAR" controls for time-specific effects. "GSP" is the lag of the gross state product in each state. "UNIT" is the lag of the ratio of total number of unit banks to total banks in each state. The first two columns of results are for the legal branch status control and the third and fourth columns are for the extent of branching control specification. Regression 1 contains a control for the regulatory environment and market structure, Regression 2 contains controls from (1) and adds a control for the state economy. t-value in parenthesis. *significant at 1% level. **significant at 5% level.***significant at 10% level. For data sources, see Appendix A.

In each era, the COMPETITION model was regressed first without the state control and then with it (regression 1 and regression 2 in tables 3 through 6). While the state control was sometimes statistically significant, it does not appear to contribute much to the overall explanatory power of the regression models.

Table 6 contains the regression results utilizing a different branch specification for the modern era. In this most recent era, it is known with more precision when both intrastate and interstate branching laws were altered. This information allows for the creation of dummy variables that reflect the legal changes in branch laws. The results indicate that allowing for both types of branching increase merger activity but only the interstate control is statistically significant. It is possible that by the 1980s and 1990s, banks had exhausted much of branching through mergers within state but were still merging across state lines since interstate banking typically came later than intrastate banking.

Table 6 Estimation Results for COMPETITION in Modern Era with Alternative Branch Specification

Variable Name	Legal Branch Status	
	Regression 1	Regression 2
INTRASTATE	0.684 (0.931)	0.369 (0.82)
INTERSTATE	3.932** (1.64)	2.728** (1.69)
NUMBER(-1)	-0.004** (1.91)	-0.003* (2.73)
YEAR	0.065 (0.723)	0.684*** (1.62)
GSP(-1)		-0.0001* (2.72)
Observations Cross-section	50	50

Table 6 Estimation Results for COMPETITION in Modern Era with Alternative Branch Specification		
Variable Name	Legal Branch Status	
	Regression 1	Regression 2
Time-series	450	450
R^2	.596	.603
F-statistic	10.792	10.901
DW Statistic	2.059	2.081

Dependent Variable: number of mergers. Independent variables: "INTRASTATE" is 0 for years of no intrastate branching, 1 otherwise for each state. "INTERSTATE" is 0 for years of no interstate branching, 1 otherwise for each state. "NUMBER" is the lag of the total number of banks in each state. "YEAR" controls for time-specific effects. "GSP" is the lag of the gross state product in each state. "UNIT" is the lag of the ratio of total number of unit banks to total banks in each state. The first two columns of results are for the legal branch status control and the third and fourth columns are for the extent of branching control specification. Regression 1 contains a control for the regulatory environment and market structure, Regression 2 contains controls from (1) and adds a control for the state economy. t-value in parenthesis. *significant at 1% level. **significant at 5% level.***significant at 10% level. For data sources, see Appendix A.

Tables 7, 8, and 9 contain the regression results for the FAILURE model for the Great Depression era and modern era respectively. Because the author does not have access to state level failure data during the national banking era, this model cannot be estimated for the earliest period. While the results of the COMPETITION model provide support for the small bankers' position against the expansion of branching, the FAILURE model provides confirmation of the large banker position that branching stabilizes banking. More specifically, all specifications of branching are of the predicted sign and indicate that more branching opportunities decrease bank failures. It is interesting, however, that the branch specifications are not consistently significant in the Great Depression era. It may be that during the depression, the impact of the wider economy mattered more than the branching structure in explaining bank failures. Given the severity of the economic downturn in this era, this would not be too surprising. Regardless, the results of the FAILURE model clearly indicate that the large banker was correct on the branch issue: branching reduces bank failures.

Table 7 Estimation Results for FAILURE in Great Depression Era				
Variable Name	Number of Banks Operating Branches		Ratio of Unit Banks to Total Banks	
	Regression 1	Regression 2	Regression 1	Regression 2
UNIT(-1)			13.131*** (1.64)	13.633 (1.45)
BRANCH(-1)	-0.278 (1.42)	-0.263 (1.39)		

Table 7 Estimation Results for FAILURE in Great Depression Era

Variable Name	Number of Banks Operating Branches		Ratio of Unit Banks to Total Banks	
	Regression 1	Regression 2	Regression 1	Regression 2
CAPASSET(-1)	-19.93 (0.97)	-18.671 (1.04)	-18.512 (1.04)	-16.97 (0.96)
LOANASSET(-1)	76.277* (2.39)	73.528** (2.24)	76.287* (2.41)	72.940* (2.24)
DEPASSET(-1)	-38.260*** (1.59)	-37.046 (1.47)	-38.882*** (1.61)	-37.431*** (1.58)
DIVERSITY(-1)	-6.555 (1.32)	-30.851 (1.32)	-11.296 (0.95)	-16.726 (0.70)
YEAR	2.255** (1.97)	2.127*** (1.76)	2.395** (2.09)	2.240** (1.87)
%BUSFAILURE(-1)	20.082* (2.63)	19.676* (2.54)	20.345* (2.67)	19.873* (2.58)
CRISIS	61.765* (9.41)	62.081* (9.37)	61.613* (9.40)	61.998 (9.39)
NUMBER(-1)		0.002** (2.21)		0.001** (1.93)
Observations				
Cross-sectional	47	47	47	47
Time-series	423	423	423	423
R^2	.536	.593	.526	.532
F	6.391	6.419	6.561	6.641
DW Statistic	2.151	2.151	2.149	2.151

Dependent Variable: Number of bank failures in each state. Independent Variables: "UNIT" is the lag of the ratio of number of unit banks to total number of banks in each state. "BRANCH" is the lag of the ratio of the number of banks operating branches to total number of banks in each state. "CAPASSET" is the lag of the ratio of total bank capital to total bank assets in each state. "LOANASSET" is the lag of the ratio of total bank loans to total bank assets in each state. "DEPASSET" is the lag of the ratio of total bank deposits to total bank assets in each state. "DIVERSITY" is the lag of the ratio of total number of bank branches to total number of banks in each state. "YEAR" controls for time-specific effects. "%BUSFAILURE" is the lag of the percent of commercial business failures in each state. "CRISIS" equals one in 1933, otherwise 0. "NUMBER" is the lag of the total number of banks in each state. Regression 1 contains a control for the regulatory environment, bank financial health controls, and the state economy. Regression 2 contains controls from (1) and adds a control for the market structure. t-value in parenthesis. *significant at 1% level. **significant at 5% level.***significant at 10% level. For data sources, see Appendix A.

Four other observations from tables 7 through 9 are worth noting. First is that DIVERSITY is not statistically significant in either era or in any model specification. This suggests that, at least

for this particular measure, diversification of bank balance sheets may have been less of a contributor to bank stability than perhaps the increased competition from branching. That is, the branching may have weeded out the weaker banks leaving a stronger system in place. This finding is consistent with the research of Carlson and Mitchener (2006) who find that, during the Great Depression, branching was more important to bank stability than diversification. Second, in all FAILURE regressions, the control for the health of the state economy was an important factor in explaining bank failures. The health of the economy clearly contributes to the health of the commercial banking sector. Third, the financial health of the bank affects bank failures. However, in the Great Depression era, the failures are explained more by the cost of deposits and the reliance on loans whereas in the modern era, the level of bank capital is statistically more valuable in explaining the failures. The fourth and final observation is that the ratio of total deposits to total assets, DEPASSET, is of the expected sign in the Great Depression but not in the Modern Era. Recall that the argument was that deposits are the cheapest source of funds and so could provide a cost advantage to banks with a higher ratio. The results in tables 8 and 9 indicate the opposite during the modern era. It may be that liability competition in the most recent era, from both commercial banks and other financial institutions, has eroded the cost advantage of deposits.

Table 8: Estimation Results for FAILURE in Modern Bank Era

Variable Name	Number of Banks Operating Branches		Ratio of Unit Banks to Total Banks	
	Regression 1	Regression 2	Regression 1	Regression 2
UNIT(-1)			6.651*** (1.69)	6.888*** (1.72)
BRANCH(-1)	-1.490** (2.29)	-2.931*** (2.02)		
CAPASSET(-1)	-4.696** (2.11)	-9.495** (2.23)	-6.809*** (1.73)	-11.561** (1.96)
LOANASSET(-1)	0.594 (0.94)	0.966 (1.02)	0.296 (0.741)	0.257 (0.82)
DEPASSET(-1)	0.004 (0.38)	0.071 (0.59)	0.0002 (0.24)	0.035 (0.31)
DIVERSITY(-1)	-0.493 (0.73)	-0.388 (0.94)	-0.328 (0.81)	-0.152 (0.75)
YEAR	1.739*** (1.98)	2.037** (2.45)	1.411** (2.19)	1.562** (2.34)
GSP(-1)	-0.0002* (4.52)	-0.0001* (4.72)	-0.0003* (5.25)	-0.0002* (5.15)
NUMBER(-1)		0.0001*** (1.99)		0.0001*** (1.94)

Table 8: Estimation Results for FAILURE in Modern Bank Era				
Variable Name	Number of Banks Operating Branches		Ratio of Unit Banks to Total Banks	
	Regression 1	Regression 2	Regression 1	Regression 2
Observations				
Cross-sectional	50	50	50	50
Time-series	450	450	450	450
R^2	.748	.766	.779	.779
F-statistic	22.384	23.847	24.294	23.880
DW statistic	1.903	1.956	1.921	1.921

Dependent Variable: Number of bank failures in each state. Independent Variables: "UNIT" is the lag of the ratio of number of unit banks to total number of banks in each state. "BRANCH" is the lag of the ratio of the number of banks operating branches to total number of banks in each state. "CAPASSET" is the lag of the ratio of total bank capital to total bank assets in each state. "LOANASSET" is the lag of the ratio of total bank loans to total bank assets in each state. "DEPASSET" is the lag of the ratio of total bank deposits to total bank assets in each state. "DIVERSITY" is the lag of the ratio of total number of bank branches to total number of banks in each state. "YEAR" controls for time-specific effects. "GSP" is the lag of the gross state product. "NUMBER" is the lag of the total number of banks in each state. Regression 1 contains a control for the regulatory environment, bank financial health controls, and the state economy. Regression 2 contains controls from (1) and adds a control for the market structure. t-value in parenthesis. *significant at 1% level. **significant at 5% level.***significant at 10% level. For data sources, see Appendix A.

Table 9: Estimation Results for FAILURE in Modern Bank Era with Alternative Branch Specification		
Variable Name	Legal Branch Status	
	Regression 1	Regression 2
INTERSTATE	-0.319 (0.94)	-0.325 (1.10)
INTRASTATE	3.093* (2.59)	3.136* (2.63)
CAPASSET(-1)	-8.090*** (1.59)	-6.856*** (1.65)
LOANASSET(-1)	0.298 (0.46)	0.245 (0.31)
DEPASSET(-1)	0.018 (0.51)	0.019 (0.53)
DIVERSITY(-1)	-0.292 (0.93)	-0.105 (0.91)
YEAR	1.252** (1.83)	1.426** (2.01)

Table 9: Estimation Results for FAILURE in Modern Bank Era with Alternative Branch Specification		
Variable Name	Legal Branch Status	
	Regression 1	Regression 2
GSP(-1)	-0.0003* (5.71)	-0.0003* (5.63)
NUMBER(-1)		0.0003** (2.04)
Observations Cross-sectional Time-series	 50 450	 50 450
R^2 F-statistic DW statistic	.782 24.310 1.941	.784 23.922 1.973

Dependent Variable: Number of bank failures in each state. Independent Variables: "UNIT" is the lag of the ratio of number of unit banks to total number of banks in each state. "INTRASTATE" is 0 for years of no intrastate branching, 1 otherwise for each state. "INTERSTATE" is 0 for years of no interstate branching, 1 otherwise for each state. "CAPASSET" is the lag of the ratio of total bank capital to total bank assets in each state. "LOANASSET" is the lag of the ratio of total bank loans to total bank assets in each state. "DEPASSET" is the lag of the ratio of total bank deposits to total bank assets in each state. "DIVERSITY" is the lag of the ratio of total number of bank branches to total number of banks in each state. "YEAR" controls for time-specific effects. "GSP" is the lag of the gross state product. "NUMBER" is the lag of the total number of banks in each state. Regression 1 contains a control for the regulatory environment, bank financial health controls, and the state economy. Regression 2 contains controls from (1) and adds a control for the market structure. t-value in parenthesis. *significant at 1% level. **significant at 5% level.***significant at 10% level. For data sources, see Appendix A.

CONCLUSIONS

This paper carefully identifies the small and large banker position on the issue of branch banking throughout U.S. history. Once identified, this research attempts to empirically test the positions in three distinct historical eras. The findings indicate that both the small banker and large banker had valid arguments. Specifically, the estimation results confirm the small bankers' position that branching leads to market concentration and reduced competition. While the small banker often spoke of monopolies, this paper does not attempt to establish the degree of market concentration but, rather, the change in the level of competition as measured by the number of bank mergers. Certainly, the threat of a pure monopoly outcome has not been validated. Rather, across all indentified bank eras, the results indicate that branching increases merger activity. Despite these results, it is important to note that merging need not increase monopoly power, especially if combined with an expansion of branches. Indeed, it is possible that the combination of more mergers and more branches could result in a more competitive environment.

There is also evidence that the large banker was right. The large banker consistently argued that branching would allow banks to become more stable and hence less prone to failure. The estimation results in this paper validate this position. However, the results within this paper also indicate that for a particular measure of diversity it seems that bank failures are not explained by a banks' ability to diversify. Previous research by Carlson and Mitchener (2006) analyze this question using data from the Great Depression and they find evidence that branching contributed quantitatively to the stability of banking rather than geographic diversity. In the end, it appears that the large banker was correct that branching would enhance stability but perhaps because it would weed out the smaller and weaker banks and not because it would improve the diversification of the balance sheet.

These findings paint a picture of a banking system that has fewer banks but more stability. A glimpse at the current market structure is consistent with this picture. Since the 1994 regulatory change that allowed for interstate banking across the country, the number of banks has steadily fallen. Between 1994 and 2007, the number of commercial banks has decreased just over 30% (fdic.gov). However, the number of branches during the same period has increased just over 43%. At the same time, the banking sector has enjoyed a period of stability as the number of bank failures has averaged 4.25 failures per year in the post interstate branching era (through 2007). However, it appears that the commercial banking sector, and wider financial sector, is sure to undergo significant change from this point forward given the turmoil of 2007 and 2008 and ongoing instability in 2009.

REFERENCES

All-Bank Statistics: U.S. 1896-1955 (1959).

Annual Report (various years). Office of the Comptroller of Currency, Washington, D.C.

Amos, Orley M. (1992). The Regional Distribution of Bank Closings in the United States from 1982 to 1988. *Southern Economic Journal*, 58(3), 805-815.

Bacon, Kenneth H. (1994, February 25). Nationwide-Bank Bill Picks Up Steam As Even Opponents See Measure Passing, *Wall Street Journal*, A3.

Banking and Monetary Statistics, 1914-1941 (1976). Washington, D.C.: Board of Governors of the Federal Reserve.

Barth, James R., Gerard Caprio, Ross Levine (2004). Bank Regulation and Supervision: What Works Best? *Journal of Financial Intermediation*, 13(2), 205-248.

Becker, Gary S. (1983). A Theory of Competition Among Pressure Groups for Political Influence. *The Quarterly Journal of Economics,* XCVIII(3), 371-400.

Berger, Allen, Anil Kashyup, and Joeseph Scalise (1995). The Transformation of the U.S. Banking Industry: What a Long, Strange Trip It's Been. *Brookings Papers on Economic Activity*, 55-218.

Bodenhorn, Howard (2003). *State Banking in Early America: A New Economic History*. New York: Oxford University Press.

Boeckle, R.M. (1937). Expansion of Branch Banking. *Editorial Research Reports*, Washington, DC: CQ Press.

Bradford, Frederick, A. (1930). Proposed National Banking Legislation. *The American Economic Review,* 20(3), 400-425.

Calomiris, Charles W. (1990). Is Deposit Insurance Necessary? A Historical Perspective. *The Journal of Economic History*, 50(2), 283-295.

Calomiris, Charles W. (2000). *U.S. Bank Deregulation in Historical Perspective*. Cambridge: Cambridge University Press.

Carlson, Mark (2004). Are Branch Banks Better Survivors? Evidence From The Depression Era. *Economic Inquiry*, 42(1), 111-126.

Carlson, Mark, and Kris James Mitchener (2006). Branch Banking, Bank Competition, and Financial Stability. *Journal of Money, Credit, and Banking*, 38(5), 1293-1328.

Cebula, Richard J. (1994). The Regional Distribution of Bank Closings in the United States: An Extension of the Amos Analysis. *Southern Economics Journal*, 61(1), 202-208.

Chapman, John M. (1934). *Concentration in Banking*. New York: Columbia University Press.

Chapman John M. and Ray B. Westerfield (1942). *Branch Banking: Its Historical and Theoretical Position in America and Abroad*. New York: Harper & Brothers Publishers.

Cochero, Steve (1991, April). Banking Reviews the Blueprints. *ABA Banking Journal*, 83(4), 33-37.

Cochero, Steve (1993, January). What Do you Get For $50 Million? Zip. *ABA Banking Journal*, 85(1), 33-36.

Collins, Charles Wallace (1926). *The Branch Banking Question*. New York: The Macmillan Company.

Congressional Globe. (1836). The Deposit Bill. Senate, 24th Congress, 1st Session, 500.

Congressional Quarterly Weekly Online. (various years), from http://library.cqpress.com/cqweekly.

Congressional Record. (various years).

Cranford, John R. (1992, January). Marketplace Forces Hold Key To Future Bank Legislation. *CQ Weekly Online*, 2, 9-14. Retrieved from http://library.cqpress.com/cqweekly.

Federal Reserve Bulletin. (various years).

Federal Reserve (1943). *Banking and Monetary Statistics, 1914-1941*, Washington, DC: Board of Governors of the Federal Reserve System.

Giedeman, Daniel C. (2005, March). Branch Banking Restrictions and Finance Constraints in Early-Twentieth-Century America. *The Journal of Economic History*, 65(1), 129-151.

Goldenweiser, E.A. (1931). Federal Reserve Committee on Branch, Group and Chain Banking: Statistical Data. Unpublished Manuscript, Federal Reserve System.

Goldenweiser, E.A. (1933) Bank Suspensions in the United States, 1892-1931. Unpublished Manuscript, Federal Reserve System.

Hammond, Bray (1957). *Banks and Politics in America, from the Revolution to the Civil War*. Princeton: Princeton University Press.

Heclo, Hugh (1978). Issue Networks and the Executive Establishment. In Anthony King (Ed.), *The New American Political System*, Washington DC: American Enterprise Institute for Public Policy Research.

Jay, Pierre (1933, January). The Structure of the Banking System. *Proceedings of the Academy of Political Science,* 15(2), 11-21.

Jayaratne, Jith and Philip E. Strahan (1999). The Benefits of Branching Deregulation. *Regulation*, 22(1), 8-16.

Kane, Edward J. (1996). De Jure Interstate Banking: Why Only Now? *Journal of Money, Credit and Banking*, 28(2), 141-161.

Kaufman, George S. (1996). Bank Failures, Systemic Risk, and Bank Regulation. *Cato Journal*, 16(1), 17-45.

Kroszner, Randall S., and Philip E. Strahan (1999, November). What Drives Deregulation? Economics and Politics of the Relaxation of Bank Branching Restrictions. *The Quarterly Journal of Economics*, 114(4), 1437-1467.

LaWare, John P. (1993). Statements to the Congress. *Federal Reserve Bulletin*, 772-780.

Levin, Ross, Gerard Caprio, James R. Barth (1999). Banking Systems Around the Globe: Do Regulation and Ownership Affect Performance and Stability? *Research Working Paper Series*, The World Bank.

Loucks, Christine (1994). The Regional Distribution of Bank Closings in the United States from 1982 to 1988: A Brief Note. *Southern Economic Journal*, 61(1), 191-194.

McChesney, Fred S. (1997). *Money for Nothing: Politicians, Rent Extraction, and Political Extortion*. Cambridge: Harvard University Press.

Newfang, Oscar (1901, December 30). Feasibility of Branch Banks. *New York Times*, D9.

New York Times. (various years).

Preston, H.H. (1924, September). Recent Developments in Branch Banking. *American Economic Review*, 443-462.

Ramirez, Carlos D. (2003). Did Branch Banking Restrictions Increase Bank Failures? Evidence from Virginia and West Virginia in the Late 1920s. *Journal of Economics and Business*, 55(4), 331-352.

Shiers, Alden F. (2002). Branch Banking, Economic Diversity and Bank Risk. *The Quarterly Review of Economics and Finance*, 42, 587-598.

Smith, Daniel R. (1994). After Interstate: 8 banks or 8,000? *ABA Banking Journal*, 86(9), 17.

Southworth, S. (1928). *Branch Banking in the United States*, New York: McGraw Hill.

Sprague, O.M.W. (1903, February). Branch Banking in the United States. *Quarterly Journal of Economics,* 17(2), 242-260.

Stigler, George J. (1971). The Theory of Economic Regulation. *Bell Journal of Economics and Management Science,* 2, 3-21.

Stiroh, Kevin J., and Philip E. Strahan (2003, October). Competitive Dynamics of Deregulation: Evidence from U.S. Banking. *Journal of Money, Credit, and Banking*, 35(5), 801-828.

U.S. Comptroller of the Currency (1876). *Annual Report of the Comptroller of the Currency*. Government Printing Office, Washington, D.C.

United States Department of Commerce (1976). *Historical Statistics of the United States Colonial Times to 1970*. Part 2.

Wall Street Journal. (various years).

Wernette, John (1932). Branch Banking in California and Bank Failures. *Quarterly Journal of Economics*, 46(2), 362-65.

Wheelock, David C. (1992, December). Regulation and Bank Failures: New Evidence From the Agricultural Collapse of the 1920s. *The Journal of Economic History*, 52(4), 806-825.

Wheelock, David C. (1995, March-April). Regulation, Market Structure, and the Bank Failures of the Great Depression. *Review*, Federal Reserve Bank of St. Louis, 27-38.

White, Eugene Nelson (1982, March). The Political Economy of Banking Regulation, 1864-1933. *The Journal of Economic History,* 42(1), 33-40.

White, Eugene Nelson (1983). *The Regulation and Reform of the American Banking System, 1900-192*9. Princeton, New Jersey: Princeton University Press.

White, Eugene Nelson (1985). The Merger Movement in Banking, 1919-1933. *Journal of Economic History*, 45(2), 285-291.

Wooldridge, Jeffrey M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press.

Appendix A: Data Sources and Description			
Data	National Bank Era[a]	Great Depression Era[b]	Modern Era
Gross State Product	NA	NA	Survey of Current Business, various years
Commercial failures	Historical Statistics of the U.S. Colonial Times to 1970, each year 1900-1910.	Historical Statistics of the U.S. Colonial Times to 1970, each year 1924-1934.	Not required.
Total Number Commercial Banks	All Bank Statistics 1896-1955, each year 1900-1910	All Bank Statistics 1896-1955, each year 1924-1934	FDIC Historical Statistics on Banking, each year 1984-1994
Total Number Unit Banks	Calculated as Total Number of Banks minus Number of Banks Operating Branches.	Calculated as Total Number of Banks minus Number of Banks Operating Branches.	FDIC Historical Statistics on Banking, each year 1984-1994
Number of Banks Operating Branches	Banking and Monetary Statistics, 1914-1941, for years 1900 and 1910. Remaining years extrapolated from regional data in Goldenweiser (1931)	Banking and Monetary Statistics, 1914-1941, for years 1925 and 1930. Goldenweiser (1931) for 1931. Remaining years extrapolated from regional data in Goldenweiser (1931)	FDIC Historical Statistics on Banking, each year 1984-1994
Number of Branches	Banking and Monetary Statistics, 1914-1941, for years 1900 and 1910. Remaining years extrapolated from regional data in Goldenweiser (1931)	Banking and Monetary Statistics, 1914-1941, for years 1925 and 1930. Goldenweiser (1931) for 1931. Remaining years extrapolated from regional data in Goldenweiser (1931)	FDIC Historical Statistics on Banking, each year 1984-1994
Number of Mergers	Calomiris (2000) provides the number of mergers and consolidations for each state for 1900-1909. 1910 data extrapolated.	Federal Reserve Bulletin, 1937, each year 1924-1934	FDIC Historical Statistics on Banking, each year 1984-1994
Total Bank Capital	All Bank Statistics 1896-1955, each year 1900-1910	All Bank Statistics 1896-1955, each year 1924-1934	FDIC Historical Statistics on Banking, each year 1984-1994
Total Bank Assets	All Bank Statistics 1896-1955, each year 1900-1910	All Bank Statistics 1896-1955, each year 1924-1934	FDIC Historical Statistics on Banking, each year 1984-1994

Appendix A: Data Sources and Description			
Data	National Bank Era[a]	Great Depression Era[b]	Modern Era
Total Bank Loans	All Bank Statistics 1896-1955, each year 1900-1910	All Bank Statistics 1896-1955, each year 1924-1934	FDIC Historical Statistics on Banking, each year 1984-1994
Total Deposits	All Bank Statistics 1896-1955, each year 1900-1910	All Bank Statistics 1896-1955, each year 1924-1934	FDIC Historical Statistics on Banking, each year 1984-1994
Number of Bank Failures	Annual Report of the Comptroller of the Currency (1910, 212-23) for national banks and Barnett (1911, 186-190) for state banks	Banking and Monetary Statistics, 1914-1941, each year 1924-1934	FDIC Historical Statistics on Banking, each year 1984-1994

Notes: [a]The national bank era sample does not include Alaska, Arizona, New Mexico, or Hawaii since these states were not yet legally formed. Further, Oklahoma has been omitted since its statehood was declared in 1907. [b]The Great Depression era does not include Alaska or Hawaii since these states were not yet legally formed.