

RISE | RICE INITIATIVE *for the* STUDY *of* ECONOMICS

RISE Working Paper 15-006

“Banking Crises, Early Warning Models,
and Efficiency”

by Pavlos Almanidis and Robin C. Sickles



RICE

Department of Economics

Baker Hall, MS22

6100 Main Street, Houston, Texas 77005

<https://economics.rice.edu>

Banking Crises, Early Warning Models, and Efficiency*

Pavlos Almanidis
International Transfer Pricing Services
Ernst & Young LLP

Robin C. Sickles
Department of Economics
Rice University

June 12, 2015

Abstract

This paper proposes a general model that combines the Mixture Hazard Model with the Stochastic Frontier Model for the purposes of investigating the main determinants of the failures and performances of a panel of U.S. commercial banks during the financial crisis that began in 2007. The combined model provides measures of the probability and time to failure conditional on a bank's performance and vice versa. Both continuous-time and discrete-time specifications of the model are considered in the paper. The estimation is carried out via the expectation-maximization algorithm due to incomplete information regarding the identity of at-risk banks. In- and out-of-sample predictive accuracy of the proposed models is investigated in order to assess their potential to serve as early warning tools.

JEL classification codes: C33, C41, C51, D24, G01, G21.

Key words and phrases: Financial distress, panel data, bank failures, semiparametric mixture hazard model, discrete-time mixture hazard model, bank efficiency.

*We would like to thank Mahmoud El-Gamal, John Bryant, Natalia Sizova, Qi Li, and Bryan Brown for their insightful comments and suggestions. We also thank seminar participants at the Economic Workshops and Seminars at Rice, Texas A&M University and the University of Queensland for their helpful suggestions and criticisms. We would like to thank Robert Adams at the Board of Governors of the Federal Reserve System for his valuable help in data construction. The views expressed in this paper are solely the responsibility of the first author and are independent from views expressed by Ernst & Young, LLP. Any errors are our own.

1 Introduction

In light of the recent 2007-2011 financial meltdown in the United States (U.S.), during which more than 400 banks and thrifts failed, that is were forced into closure by regulatory agencies¹, the need for effective regulations and intervention policy that would identify and resolve future crises and undertake prompt corrective actions to resolve such crises with minimal cost in a timely fashion has been recognized as essential to the health of the U.S. economy. In this paper we only consider failed banks as ones that appear on the FDIC's failed bank list and ceased their operation due to reasons other than merger or voluntary liquidation, or that remained inactive or no longer were regulated by the Federal Reserve. The 2007-2011 financial crisis, which originated from the secondary market for residential mortgage-backed securities (RMBS) immediately after the collapse of the housing bubble in 2006, caused severe losses to banks and in particular large banks, which were highly involved in the RMBS market. At the same time and as a result of the large banks' widespread distress and contagion effects, the number of problem banks on the watch list maintained by the Federal Deposit Insurance Corporation (FDIC) dramatically increased. Systemically important financial institutions at risk, commonly described as too-big-to-fail, received heavy doses of government funds through the Troubled Asset Relief Program (TARP) from regulatory authorities who apparently believed that the banks' failures would impose greater systemic risk that could substantially damage the economy and lead to conditions similar to, or possibly exceeding, those of the Great Depression. The financial crisis footprint was not the same across the states. Those that experienced the most failures were California, Florida, Georgia and Illinois, accounting for more than half of the failures in the U.S.

Banking crises are not a new phenomena in the U.S. economy² and regulatory authorities have always considered banking failures as a major public policy concern, because of the special role that banks play in the economic network and in the implementation of an effective monetary policy. The distinguishing characteristic of the banking crisis of 2007-2011 from those in the 1980s and 1990s, however, is that failures were not limited to small financial institutions. Rapid credit expansion and low quality loans and investments made during a period of economic expansion mainly took their toll on large multi-billion dollar financial institutions. Approximately one in five banks that failed had assets of over \$1 billion and in 2008 thirty-six percent were large banks, among them the largest bank failure in the history of the U.S., that of Washington Mutual with \$307 billion in assets.³ That same year saw Lehman Brothers file for Chapter 11 bankruptcy protection and IndyMac Bank with \$32 billion in assets taken over by the FDIC.⁴ These large financial institution

¹According to the Federal Deposit Insurance Corporation's Failed Bank List available at: <http://www.fdic.gov/bank/individual/failed/banklist.html>

²The Great Depression of 1930s and savings and loan (S&L) crisis of the 1980s and 1990s are the two most obvious examples from the last century.

³Continental Illinois Bank and Trust Company of Chicago failed in 1984 and had one-seventh of Washington Mutual's assets.

⁴Chapter 11 permits reorganization under the bankruptcy laws of the United States. A financial institution filing for Chapter 11 bankruptcy protection usually proposes a plan of reorganization to keep its business alive and pay its creditors over time.

failures created large uncertainties about the exposure of other financial institutions (healthy and troubled) to additional risks, reduced the availability of credit from investors to banks, drained the capital and money markets of confidence and liquidity, triggered the failure of smaller community banks⁵, and raised fears of severe instability in the financial system and the global economy.

In the U.S., the FDIC and state banking regulatory authorities are responsible for the identification and resolution of insolvent institutions. A bank is considered at a risk of immediate closure if it is unable to fulfil its financial obligations the next day or its capital reserves fall below the required regulatory minimum.⁶ The FDIC is required to resolve outstanding issues with problem banks in a manner that imposes the least cost on the deposit insurance fund (DIF) and ultimately on the taxpayer. Thus, early detection of insolvent institutions is of vital importance, especially if the failure of those institutions would pose a serious systemic risk on the financial system and the economy as a whole. The FDIC and state authorities utilize on-site and off-site examination methods in order to determine which institutions are insolvent and, thus, should be either closed or be provided financial assistance in order to rescue them. The off-site examinations are typically based on statistical and other mathematical methods and constitute complementary tools to the on-site visits made by supervisors to institutions considered at risk. There are three advantages to off-site versus on-site examinations: (i) the on-site examinations are more costly as they require the FDIC to bear the cost of visits and to retain extra staff during times when economic conditions are stable; (ii) the on-site examinations are usually time-consuming and cannot be performed with high frequency; and (iii) the off-site examinations can help allocate and coordinate the limited on-site examination resources in an efficient way with priority given to financial institutions facing the most severe challenges. The major drawback of statistically-based off-site tools is that they incorporate estimation errors which may affect the classification of banks as failure and nonfailures. An effective off-site examination tool must aim at identifying problem banks sufficiently prior to the time when a marked deterioration of their financial health would occur. Therefore, it is desirable to develop a model which would identify future failures with a high degree of accuracy in a timely manner and would rarely flag healthy banks as being at risk of closure.

This paper develops an early warning model of bank troubles and failures based on the Mixture Hazard Model (MHM) of Farewell (1977, 1982) with continuous and discrete time specifications.⁷ MHM effectively combines the static model, which is used to identify troubled banks, and the duration model, which provides estimates of the probability of failure along with the timing of closure of such troubled banks. We view the financial

⁵Community banks are banks with assets sizes of \$1 billion or less. Their operation is oftentimes limited to rural communities and small cities. Community banks usually engage in traditional banking activities and provide more personal-based services.

⁶Under the current regulations issued by the Basel Committee on Banking Supervision (Basel II&III), a bank is considered as failed if its ratio of Tier 1 (core) capital to risk-weighted assets is 2% or lower. This ratio must exceed 4% to avoid supervisory intervention and prompt corrective action as underlined in Federal Deposit Insurance Corporation Improvement Act (FDICIA) of 1992. A bank with ratio of 6% or above is considered as a well-capitalized bank.

⁷Applications of the discrete-time version of the MHM can be found in Gonzalez-Hermosillo et al.(1997), Yildirim (2008) and Topaloglu and Yildirim (2009).

crisis as a negative shock that affected banks in an unequal way. Well-capitalized, well-prepared, and prudently-managed institutions may have felt relatively little distress during the financial turmoil. On the other hand, poorly-managed banks that previously engaged in risky business practices faced an increased probability of their being on the FDIC watch list and, subsequently forced into closure or merger with a surviving bank by regulatory authorities. Unlike standard duration models, which assume that all banks are at the risk of failure, we implicitly assume that a proportion of banks will survive for a sufficiently long time after the end of a crisis and thus are incapable of entering an absorption state. In other words, we assume that the probability of failure for a bank that has never been on the watch list is arbitrarily close to zero. The MHM is appropriate for dealing with this issue as it is able to distinguish between healthy and at-risk of failure banks.

One of our (testable) assumptions concerns the fact that banks with low performance, as calculated by the radial measure of realized outcome to the maximum potential outcome, will increase their probability of failure. An inefficiently-managed bank could cumulatively save valuable funds by improving its performance. The saved funds often prove to be vital in servicing a bank's short-term obligations during financial crisis periods when interbank markets suffer from poor liquidity, and would therefore prevent the bank to need to draw on shareholders' equity. Shareholders' equity is the most expensive source of financing, the reduction of which would trigger on-site examination by regulators and possibly would place the bank on the watch list of problem banks. On-site examination subsequently would redirect the bank management's focus on clearing problem accounts rather than on improving its overall performance and thus could make it even less efficient. This process could continue in a spiral fashion, deteriorating the bank's financial assets and the capital. To account for this mutual effect, we employ a single step joint estimation procedure proposed by Tsionas and Papadogonas (2006), wherein a stochastic frontier model (SFM) is jointly estimated with a frailty model.

A challenge that we face in this paper is the incomplete information associated with the troubled banks on the watch list of the FDIC. Each quarter the FDIC releases the number of problem banks, but their identities are not publicly disclosed. To address this problem of missing information, we make an assumption that a bank that failed was on this list and based on available information we make a prediction of which banks potentially could be on this list through an expectation-maximization (EM) algorithm, which is designed to deal with this type of incomplete information. We also follow a forward step-wise procedure in model building and covariates selection, which is not only based on the conventional measures of the goodness-of-fit and statistical tests, but also on the contribution of these covariates to the predictive accuracy of the proposed models.

Finally, our model recognizes the fact that insolvency and failure are two different events. The realization of the first event is largely attributed to the actions undertaken by a bank itself, while the second usually occurs as a result of regulators' intervention following its insolvency. Supervisors typically tend not to seize an insolvent bank unless it has no realistic probability of survival and its closure does not threaten the soundness and the stability of the financial system. Based on the above considerations, we are able to assess the type I

and type II errors implicit in bank examiners' decision process when closing banks.⁸ We find that the within sample and out-of-sample average of the two misclassification errors is less than 6% and 2%, respectively, for our preferred model. We also find that the predictive power of our model is quite robust when using estimates derived from different sub-periods of the financial crisis.

The remainder of the paper is organized as follows. In Section 2 we provide a brief review of banking crisis models. Section 3 describes the potential decision rule adopted by the regulatory authorities in determining and closing insolvent banks, which naturally will lead to the MHM. Two variants of the MHM are discussed, the continuous-time semiparametric proportional MHM and discrete-time MHM. In section 4 we discuss the joint MHM-SFM. Section 5 deals with empirical specification issues and the data description. Estimation, testing, and predictive accuracy results are provided in section 6, along with a comparison of various models and specifications. Section 7 contains our main conclusions.

2 Banking Crisis Studies

Accurate statistical models that serve as early warning tools and that potentially could be used as an alternative or complement to costly on-site visits made by supervisors have been well documented in the banking literature. These models have been applied successfully to study banking and other financial institutions' failures in the U.S. and in other countries. As the literature that deals with bankruptcy prediction of financial institutions is vast and there are a myriad of papers that specifically refer to the banking industry failures, we will discuss only few the studies that are closely related to our work and are viewed as early warning models.

The more widely-used statistical models for bankruptcy prediction are the single-period static probit/logit models and the methods of discriminant analysis. These method usually estimate the probability that an entity with specific characteristics will fail or survive within a certain time interval. The timing of the failure is not provided by such models. Applications of the probit/logit models and discriminant analysis can be found in Altman (1968), Meyer and Pifer (1970), Deakin (1972), Martin (1977), Lane et al. (1986), Cole and Gunther (1995, 1998), Cole and Wu (2011), among others.

Others in this literature have employed the Cox proportional hazard model (PHM) and the discrete time hazard model (DTHM) to explain banking failures and develop early warning models.⁹ In the hazard model the dependent variable is time to the occurrence of some specific event, which can be equivalently expressed either through the probability distribution function or the hazard function, which provides the instantaneous risk of failure at some specific time conditional on the survival up to this time. The PHM has three advantages over the static probit/logit models: (i) it provides not only the measure of probability of failure (survival), but also the probable timing of failure; (ii) it accommodates

⁸Typically, a type I error is defined as the error due to classifying a failed bank as a non-failed bank, while a type II error arises from classifying a non-failed bank as a failed bank.

⁹A thorough discussion of hazard models can be found in Cox (1972), Lancaster (1990), Kalbfleisch and Prentice (2002), and Klein and Moeschberger (2003).

censored observations, those observations that survive through the end of the sample period; and (iii) it does not make strong assumptions about the distribution of duration times. The disadvantage of the PHM model is that it requires the hazard rate to be proportional to the baseline hazard between any two cross-sectional observations. Moreover, inclusion of time-varying covariates is problematic. The DTHM, on the other hand, easily allows for time-varying covariates and has the potential to provide more efficient estimates and improved predictions. The application of PHM to the study the U.S. commercial banking failures was undertaken by Lane et al. (1986), Whalen (1991), as well as in Wheelock and Wilson (1995, 2000).¹⁰

Barr and Siems (1994) and Wheelock and Wilson (1995, 2000) were the first to consider inefficiency as a potential influential factor explaining U.S. commercial banking failures during the earlier crisis. They estimated the efficiency scores with Data Envelopment Analysis (DEA) techniques, which were then used in a static model to predict banking failures.¹¹ Wheelock and Wilson (1995, 2000), on the hand, included inefficiency scores among their regressors to allow these to affect the probability of failure and acquisitions by other banks in the PHM framework. They employed three measures of radial technical inefficiency, namely the parametric cost inefficiency measure, the nonparametric input distance function measure, and the inverse of the nonparametric output distance function measure. According to the authors, the first two had a statistically significant (positive) effect on the probability of failure, while only the first measure significantly decreased the acquisition probability. The estimation of the models was conducted in two stages. The first stage involved the parametric or nonparametric estimation of inefficiency scores. In the second stage these scores were used among the explanatory variables to investigate their effect on the probabilities of failure and acquisition. Tsionas and Papadogonas (2006) criticize the two-step approach, arguing that this may entail an error-in-variables bias as well as introduce an endogenous auxiliary regressor.

3 Mixture Hazard Model

Our banking failure modelling approach is based on the rules and policies that regulatory authorities implement in order to identify problem banks that subsequently fail or survive.¹² We first let H_{it} define the financial health indicator of bank i at time t and assume that there is a threshold level of it, H_{it}^* , such that if the financial health falls bellow this level then the bank is considered at risk of closure by regulatory authorities. Second, we let the difference between H_{it}^* and H_{it} , denoted by h_{it}^* , be dependent on bank-specific financial metrics and market variables as follows:

$$h_{it}^* = H_{it}^* - H_{it} = x_{it}'\boldsymbol{\beta} + e_{it} \tag{1}$$

¹⁰Shumway (2001), Halling and Hayden (2006), Cole and Wu (2009), and Torna (2010) provide non-banking applications, along with arguments for using the DTHM over the static models and PHM.

¹¹DEA, which was proposed by Charnes et al. (1978), is a nonparametric approach that estimates a relative efficiency score for a bank based on linear programming techniques.

¹²See Kasa and Spiegel (2008) on various regulatory closure rules.

where e_{it} represents the error term, which is assumed to be identically and independently distributed (*iid*) across observations and over time.¹³

The financial health threshold of a particular bank is a composite and oftentimes subjective index and its lower bound is not observable to the econometrician; therefore, h_{it}^* is not observable as well. Instead a binary variable h_{it} can be defined such that

$$h_{it} = \begin{cases} 1 & \text{if } h_{it}^* > 0 \\ 0 & \text{if } h_{it}^* \leq 0 \end{cases}$$

Based on the above, the probability that a bank will become a *problem* bank is given by

$$P(h_{it} = 1) = P(h_{it}^* > 0) = P(e_{it} > -x'_{it}\boldsymbol{\beta}) = F_e(x'_{it}\boldsymbol{\beta})$$

where F_e is the cumulative distribution function of the random error e , which can be assumed to be either normally distributed (probit model) or logistically distributed (logit model).

Specification of the likelihood function then follows that of the standard hazard model, wherein a nonnegative random variable T represents the time to failure of a bank within a given period of time. This is characterized by the conditional probability density function f_T and the cumulative distribution function F_T . A binary variable d_i is also specified and takes on a value of 1 for observations that fail at time t and 0 for observations that are right censored (i.e., when a bank does not fail by the end of the sample period or disappears during the period for reasons other than failure).¹⁴ Assuming that the rate at which regulatory authorities tend to seize healthy banks is arbitrary close to zero, the likelihood function for a bank i is given by

$$L_i(\boldsymbol{\theta}; x, w) = [F_e(x'_i\boldsymbol{\beta})\lambda_i^p(t; w_i)S^p(t; w_i)]^{d_i} (F_e(x'_i\boldsymbol{\beta})S^p(t; w_i) + [1 - F_e(x'_i\boldsymbol{\beta})])^{1-d_i} \quad (2)$$

where S^p is a survivor function, which represents the probability that a problem bank will survive for a period longer than t and λ^p represents the hazard rate or probability that such bank will fail during the next instant, given that it was in operation up until this time. The $\boldsymbol{\theta}$ represents the parameter vector, while x and w are covariates associated with the probability of being problem and failed, respectively. A detailed derivation of the likelihood function is provided in Appendix A of this paper. After rearranging the expression in (2) and dropping the superscript from measures pertaining to problem banks to reduce notational clutter, the sample likelihood for all banks can be written as:

¹³The *iid* assumption of the error term can be relaxed in the panel data context by assuming $e_{it} = \mu_i + \xi_{it}$ with $\mu_i \sim N(0, \sigma_\mu^2)$ and $\xi_{it} \sim N(0, \sigma_\xi^2)$ independent of each other. This adds an additional complication to the model and it is not pursued in this paper.

¹⁴In this paper, failed banks are only considered as the banks that appear on the FDIC's failed bank list. Banks that ceased their operation due to reasons other than failure (e.g., merger or voluntary liquidation) or remained inactive or are no longer regulated by the Federal Reserve, have censored duration times.

$$\begin{aligned}
L(\boldsymbol{\theta}; x, w, d) &= \prod_{i=1}^n L_i(\boldsymbol{\theta}; x, w, d) \\
&= \prod_{i=1}^n F_e(x_i \boldsymbol{\beta})^{h_{it}} (1 - F_e(x_i \boldsymbol{\beta}))^{1-h_{it}} \{\lambda_i(t; w_i)\}^{d_i h_{it}} \{S_i(t; w_i)\}^{h_{it}}
\end{aligned} \tag{3}$$

If T is assumed to be a time-varying variable, then the model can be estimated based on the proportional hazards assumption (Cox, 1972), which unfortunately does not allow for time-varying covariates. Following Kuk and Chen (1992) and Sy and Taylor (2000), the survivor and hazard functions in this case are given by the expressions below:

$$\lambda_i(t; w_i) = \lambda_0(t) \exp(w_i' \boldsymbol{\alpha}) \text{ and } S_i(t; w_i) = S_0(t)^{\exp(w_i' \boldsymbol{\alpha})} \tag{4}$$

where, $\lambda_0(t)$ and $S_0(t)$ define the conditional baseline hazard function and baseline survivor function, respectively. These are nonnegative functions of time only and are assumed to be common to all banks at risk.

The discrete-time version of the model on the other hand is more flexible and adds more dynamics to the model by allowing for inclusion of time-varying covariates. This specification, however, requires that the time-varying regressors remain unchanged in the time interval $[t, t + 1]$. The survivor and hazard functions in the discrete-time MHM can be shown to be derived as:¹⁵

$$S_{ij}(t; w, u) = \left[\prod_{j=1}^{t_i} \frac{1}{1 + \exp(w'_{ij} \boldsymbol{\alpha})} \right] \text{ and } \lambda_{ij}(t; w) = 1 - \frac{S(t_{ij})}{S(t_{i,j-1})} \text{ for } j = 1, 2, \dots, t_i.$$

In what follows, we refer to the continuous-time MHM as Model I and the discrete-time MHM as Model II. Following the standard nomenclature in the medical and biological sciences, we refer to the portion of the model that assesses the financial health of a bank as the **incidence** component and the portion of the model that assesses survival times as the **latency** component.

If h_{it} is observed by the econometrician for each individual bank as it is by regulators then the estimation process reduces to that of the standard MHM. However, as discussed above h_{it} is only partially observed by the econometrician. We address this problem of incomplete information by utilizing the EM algorithm to deal with the missing data. The EM algorithm consists of two iterative steps: the expectation (E) step and the maximization (M) step. The expectation step involves the projection of an appropriate functional (likelihood or log-likelihood function) containing the augmented data on the space of the original (incomplete) data. Thus, the missing data are first estimated, given the observed data and the initial estimates of the model parameters, in the E step. In the M step the function is maximized while treating the incomplete data as known. Iterating between these two steps yields estimates that under suitable regulatory conditions converge to the maximum likelihood

¹⁵See Cox and Oakes (1984), Kalbfleisch and Prentice (2002), and Bover et al. (2002) for discussion on discrete-time proportional hazard models.

estimates (MLE).¹⁶

To implement the EM algorithm first consider the expectation of the full log-likelihood function with the respect to the h_{it} and the data, which completes the E step of the algorithm. Linearity of log-likelihood function with respect to the h_{it} considerably facilitates the calculations and the analysis.

The log-likelihood for the i^{th} observation in the M step is given by:

$$E_{h|X,W,\theta,\lambda_0}^{(M)} [L_i(\boldsymbol{\theta}; x, w, d)] = \tilde{h}_{it}^{(M)} \log [F_e(x_i\boldsymbol{\beta})] + (1 - \tilde{h}_{it}^{(M)}) \log [1 - F_e(x_i\boldsymbol{\beta})] \quad (5) \\ + \tilde{h}_{it}^{(M)} d_i \log [\lambda_i(t; w_i)] + \tilde{h}_{it}^{(M)} \log [S_i(t; w_i)]$$

where \tilde{h}_{it} is the probability that the i^{th} bank will eventually belong to the group of problem banks at time t , conditioned on the observed data and the model parameters. It represents the fractional allocation to the problem banks and is given by:

$$\tilde{h}_{it}^{(M)} = E [h_{it} | \boldsymbol{\theta}^{(M)}, Data] = \Pr(h_{it}^{(M)} = 1 | t_i > T_i) \quad (6) \\ = \begin{cases} \frac{F_e(x'_i\boldsymbol{\beta}^{(M)})S_i(t;w_i)}{F_e(x'_i\boldsymbol{\beta}^{(M)})S_i(t;w_i)+(1-F_e(x'_i\boldsymbol{\beta}^{(M)}))} & \text{if } d_i = 0 \\ 1 & \text{otherwise} \end{cases}$$

In Model I, the nuisance baseline hazard function λ_0 is estimated nonparametrically from the profile likelihood function as:

$$\hat{\lambda}_0(t) = \frac{N(t_i)}{\sum_{j \in R(t_i)} \tilde{h}_{jt} \exp(w'_j\boldsymbol{\alpha})} \quad (7)$$

where $N(t_i)$ is the number of failures and $R(t_i)$ is the set of all individuals at risk at time t_i , respectively. Substituting (7) into (5) leads to the M step log-likelihood for Model I:

$$\tilde{L}(\boldsymbol{\theta}; x, w, \tilde{h}) = \sum_{i=1}^n \left\{ \tilde{h}_{it} \log F_e(x'_i\boldsymbol{\beta}) + (1 - \tilde{h}_{it}) \log(1 - F_e(x'_i\boldsymbol{\beta})) \right\} \quad (8) \\ + \sum_{i=1}^N \left\{ w'_i\boldsymbol{\alpha} - N(t_i) \log \left(\sum_{j \in R(t_i)} \tilde{h}_{jt} \exp(w'_j\boldsymbol{\alpha}) \right) \right\} \\ = L_1(\boldsymbol{\beta}; x, \tilde{h}) + \tilde{L}_2(\boldsymbol{\alpha}; w, \tilde{h})$$

The full implementation of the EM algorithm involves the following four steps:

- Step 1: Provide an initial estimate for the parameter $\boldsymbol{\beta}$ and estimate the ordinary MHM in order to obtain the starting values for λ_0 ;
- Step 2 (E step): Compute \tilde{h}_{it} from (6) based on the current estimates and the observed

¹⁶For more discussion on the EM algorithm and its convergence properties and limitations see Dempster et al. (1977) as well as McLachlan and Krishnan (1996).

data;

- Step 3 (M step): Update the estimate of parameter β using (5); and
- Step 4: Iterate between steps 2 and 3 until convergence is reached.¹⁷

Alternatives to the EM method can also be utilized. For example, in his study of the recent U.S. commercial banking failures Torna (2010) attempted to identify troubled banks on the FDIC's watch list based on their tier 1 capital ranking. Banks were ranked according to their tier 1 capital ratio and the number of banks with the lowest value were selected to match the number provided by the FDIC in each quarter. Other ratios, such as Texas ratio, also can be utilized to deduce the problem banks. The Texas ratio was developed by Gerard Cassidy to predict banking failures in Texas and New England during recessionary periods of the 1980's and 1990's. It is defined as the ratio of nonperforming assets to total equity and loan-loss reserves and banks with ratios close to one are identified as high risk. There are at least two limitations to these approaches besides their crude approximation. First, they ignore other variables that play a pivotal role in leading banks to a distressed state. For example, ratios based on nonperforming loans are major indicators of difficulties that bank will face in near future, even if their current capital ratios are at normal levels. Second, financial ratios that are used to classify banks as healthy or troubled cannot be subsequently employed as determinants due to a possible endogeneity problem.

3.1 Combined SFM and MHM Model

In this section we consider the efficiency performance of a bank as a determinant of the probability of being both a problem bank and one that subsequently fails. The efficiency performance of a firm relative to the best practice (frontier) technology firm was formally considered by Debreu (1951) and Farrell (1957). Aigner et al. (1977), Meeusen and van den Broeck (1977), and Battese and Cora (1977) introduced the parametric stochastic frontier model (SFM). In the SFM the error term is assumed to be multiplicative in a levels specification of the production or of one of its dual presentations, such as the cost function we use in our analysis, and is composed of two parts: (i) a one-sided error term that captures the effects of inefficiencies relative to the stochastic frontier; and (ii) a two-sided error term that captures random shocks, measurement errors and other statistical noise.¹⁸

The general SFM is represented by the following functional relationship:

$$y_{it} = g(z_{it}; \boldsymbol{\eta}) \exp(\varepsilon_{it}) \quad (9)$$

where the dependent variable y_{it} could represent cost, output, profit, revenue and so forth, z_{it} is a vector of independent regressors, and $g(\cdot)$ is the frontier function, which can be

¹⁷Convergence to a stationary point in the EM algorithm is guaranteed since the algorithm aims at increasing the log-likelihood function at each iteration stage. The stationary point need not, however, be a local maximum. It is possible for the algorithm to converge to local maxima or saddle points. We check for these possibilities by selecting different starting values and checking for the proper signs of the Hessian.

¹⁸Excellent surveys of frontier models and their applications are found in Kumbhakar and Lovell (2000) and Greene (2008).

either linear or non-linear in coefficients and covariates. Depending on the particular dual representation of technology specified, $\varepsilon = v \pm u$ [$= \log y_{it} - \log g(z_{it}; \boldsymbol{\eta})$] represents the composed error term, with v_{it} representing the noise and u_i the inefficiency process. The noise term is assumed to be *iid* normally distributed with zero mean and constant variance. Inefficiencies are also assumed to be *iid* random variables with distribution function defined on the domain of positive numbers ($u \in R_+$). Both v and u are assumed to be independent of each other and independent of the regressors.¹⁹ In this paper, we follow Pitt and Lee (1981) and assume that the inefficiency process is a time-invariant random effect, which follows the half-normal distribution (i.e., $u_i \sim N^+(0, \sigma_u^2)$).

Under the above assumptions the marginal distribution of the composed error term, which for the production or cost frontier model is derived as:

$$f_\varepsilon(\varepsilon_{it}) = \frac{2}{(2\pi)^{T_i/2} \sigma_v^{T_i-1} \sigma} \exp \left[-\frac{\varepsilon'_{it} \varepsilon_{it}}{2\sigma_v^2} + \frac{\bar{\varepsilon}_i^2 \lambda^2}{2\sigma^2} \right] \left[1 - \Phi \left(\frac{T_i \bar{\varepsilon}_i \lambda}{\sigma} \right) \right] \quad (10)$$

where $\sigma = \sqrt{\sigma_v^2 + T_i \sigma_u^2}$, $\lambda = \sigma_u / \sigma_v$, and $\bar{\varepsilon}_i = (1/T_i) \sum_{t=1}^{T_i} \varepsilon_{it}$.²⁰ The parameter λ is the signal-to-noise ratio and measures the relative allocation of total variation to the inefficiency term. In practice we can use an alternative parametrization, called the γ -parameterization, which specifies $\gamma = \sigma_u^2 / \sigma^2$.²¹

It can be also shown (see Jondrow et al., 1982) that the conditional distribution of the inefficiency term is given by

$$f_{u|\varepsilon}(u_i | \varepsilon_{it}) = \frac{f_{\varepsilon, u}(\varepsilon_i, u_i)}{f_\varepsilon(\varepsilon_i)} = \frac{\frac{1}{\sigma} \phi \left(\frac{u_i - \mu_i^*}{\sigma} \right)}{\left[1 - \Phi \left(-\frac{\mu_i^*}{\sigma} \right) \right]} \quad (11)$$

where $f_{u|\varepsilon}(\cdot)$ represents the normal distribution truncated at 0 with mean $\mu_i^* = -T_i \bar{\varepsilon}_i \sigma_u^2 / \sigma^2 = -T_i \bar{\varepsilon}_i \gamma$ and variance $\sigma_*^2 = \sigma_u^2 \sigma_v^2 / \sigma^2 = \gamma \sigma^2 (1 - \gamma T_i)$, and $\phi(\cdot)$ and $\Phi(\cdot)$ are, respectively, the pdf and cdf functions of the standard normal distribution. The mean or the mode of this conditional distribution function provides an estimate of the technical inefficiency.

In the absence of any effect of the inefficiencies on the probability of being troubled and failed, (10) and (11) can be employed to obtain the maximum likelihood estimates of model parameters and efficiency scores. However, consistent and efficient parameter estimates cannot be based solely on the frontier model when there is feedback between this measure of economic frailty and the likelihood of failure and the ensuing tightening of regulatory supervision. There is a clear need for joint estimation of the system when the decision of a firm is affected by these factors.

¹⁹The assumption of independence of the inefficiency term and the regressors is restrictive, but is necessary for our current analysis. Its validity can be tested using the Hausman-Wu test. In the panel data context, this assumption can be relaxed by assuming that inefficiencies are fixed effects or random effects correlated with all or some of the regressors (Hausman and Taylor, 1981; Cornwell, Schmidt, and Sickles, 1990).

²⁰The cost frontier is obtained by reversing the sign of the composed error.

²¹This reparametrization is desirable as the γ parameter has compact support, which facilitates the numerical procedure of maximum likelihood estimation, hypothesis testing, and establishing the asymptotic normality of this parameter.

In deriving the likelihood function for this model, we maintain the assumption that censoring is non-informative and statistically independent of h_i . Following Tsionas and Papadogonas (2006) we also assume that conditional on inefficiency and the data the censoring mechanism and h_i are independent of the composed error term. To simplify notations, let $\mathbf{\Omega}_i = \{x_i, w_i, z_i\}$ denote the set of covariates and $\Theta = \{\boldsymbol{\beta}, \boldsymbol{\alpha}, \delta_1, \delta_2, \boldsymbol{\eta}, \sigma_v^2, \sigma_u^2\}$ be the vector of the structural and distributional parameters. The observed joint density of the structural model for bank i , given h_i and after integrating out the unobserved inefficiency term, can be written as:

$$\begin{aligned}
L_i(y_i, h_i, d_i | \mathbf{\Omega}_i, \Theta') &= \int_0^\infty F_e(x_i' \boldsymbol{\beta} + \delta_1 u_i)^{h_i} (1 - F_e(x_i' \boldsymbol{\beta} + \delta_1 u_i))^{1-h_i} & (12) \\
&\quad \times \{\lambda_i(t; w_i, u_i)\}^{d_i h_i} \{S_i(t; w_i, u_i)\}^{h_i} \underbrace{f_v(\varepsilon_{it} \pm u_i) f(u_i)}_{f_\varepsilon(\varepsilon) f_{u|\varepsilon}(u|\varepsilon)} du_i \\
&= f_\varepsilon(\varepsilon_{it}) \int_0^\infty F_e(x_i' \boldsymbol{\beta} + \delta_1 u_i)^{h_i} (1 - F_e(x_i' \boldsymbol{\beta} + \delta_1 u_i))^{1-h_i} \\
&\quad \times \{\lambda_i(t; w_i, u_i)\}^{d_i h_i} \{S_i(t; w_i, u_i)\}^{h_i} f_{u|\varepsilon}(u|\varepsilon) du_i.
\end{aligned}$$

The hazard rate and survival function for the continuous-time counterpart of the model are now given by:

$$\lambda_i(t; w_i, u_i) = \lambda_0(t) \exp(w_i' \boldsymbol{\alpha} + \delta_2 u_i) \text{ and } S(t; w_i) = S_0(t)^{\exp(w_i' \boldsymbol{\alpha} + \delta_2 u_i)}$$

It should be noted that the above model is rather a general one and consists of three individual parts: (1) the SFM; (2) the probit/logit model for the incidence part; and (3) the standard hazard model for the latency part. Each of these three models are nested within the general model. For example, if there is no association between inefficiency and the probability of being troubled or failed ($\delta_1 = 0$ and $\delta_2 = 0$), then (12) consists of two distinct parts, the SFM and the MHM. Both can be estimated separately using the methods outlined in the previous sections.

The integral in the joint likelihood (12) has no closed form solution and thus the maximization of this function requires numerical techniques, such as simulated maximum likelihood (SML) or Gaussian quadrature.²² In SML the sample of draws from $f_{u|\varepsilon}(\cdot)$ are required to approximate the integral by its numerical average (expectation). As such, the

²²Tsionas and Papadogonas (2006) employed the Gaussian quadrature in estimation of the model where the technical inefficiency has a potential effect on firm exit. Sickles and Taubman (1986) used similar methods in specifying structural models of latent health and retirement status, while controlling for multivariate unobserved individual heterogeneity in the retirement decision and in morbidity.

simulated log-likelihood function for the i^{th} observation becomes:

$$\begin{aligned}
L_i = \log L_i(y_i, h_i, d_i | \boldsymbol{\Omega}_i, \boldsymbol{\Theta}') &= Constant - \frac{(T_i - 1)}{2} \log \sigma^2 (1 - \gamma T_i) \\
&- \frac{1}{2} \log \sigma^2 + \log \left[1 - \Phi \left(\frac{T_i \bar{\varepsilon}_i \sqrt{\gamma / (1 - \gamma)}}{\sigma} \right) \right] - \frac{\varepsilon'_{it} \varepsilon_{it}}{2\sigma^2(1 - \gamma T_i)} + \frac{\bar{\varepsilon}_i^2 \gamma}{2\sigma^2(1 - \gamma)} \\
&+ \log \frac{1}{J} \sum_{j=1}^J \left\{ F_e(x_i \boldsymbol{\beta} + \delta_1 u_{ij})^{h_{it}} [1 - F_e(x_i \boldsymbol{\beta} + \delta_1 u_{ij})]^{1-h_{it}} [\lambda_i(t; w_i, u_{ij})]^{d_i h_{it}} [S_i(t; w_i, u_{ij})]^{h_{it}} \right\}
\end{aligned} \tag{13}$$

where u_{is} is a random draw from the truncated normal distribution $f_{u|\varepsilon}(\cdot)$ and J is the number of draws. We utilize the inverse cdf method to efficiently obtain draws from this distribution as:

$$u_{ij} = \mu_i^* + \sigma_* \Phi^{-1} \left[U_{is} + (1 - U_{is}) \Phi \left(-\frac{\mu_i^*}{\sigma_*} \right) \right] \tag{14}$$

where U is a random draw from uniform $U[0, 1]$ distribution or a Halton draw. By substituting (14) into (13) and treating the h_{it} s as known we can maximize the log-likelihood function $L = \sum_i L_i$ by employing standard optimization techniques and obtain the model parameters.

Finally, after estimating the model parameters, the efficiency scores are obtained as the expected values of the conditional distribution, in the spirit of Jondrow et al. (1982):

$$\hat{u}_i = E \left[u_i | \hat{\varepsilon}_i, \tilde{h}_i, d_i, \boldsymbol{\Omega}_i, \boldsymbol{\Theta}' \right] = \frac{\int_0^\infty u_i G(u_i; \boldsymbol{\Theta}) f_{u|\varepsilon}(u|\varepsilon) du_i}{\int_0^\infty G(u_i; \boldsymbol{\Theta}) f_{u|\varepsilon}(u|\varepsilon) du_i} \tag{15}$$

where $G(u_i; \boldsymbol{\Theta}) = \tilde{F}(x'_i \boldsymbol{\beta} + \delta_1 u_i)^{\tilde{h}_{it}} \left[1 - \tilde{F}(x'_i \boldsymbol{\beta} + \delta_1 u_i) \right]^{1-\tilde{h}_{it}} [\lambda_i(t; w_i, u_i)]^{d_i \tilde{h}_{it}} [S(t; w_i, u_i)]^{\tilde{h}_{it}}$. The integrals in the numerator and denominator are calculated numerically by the SML method and the efficiency score of i^{th} firm is estimated as $TE_i = \exp(-\hat{u}_i)$. It is straightforward to check that if δ is zero then (15) collapses to the formula derived by Jondrow et al. for production frontiers (i.e., $\hat{u}_i = E[u_i | \hat{\varepsilon}_i] = \mu_* + \sigma \phi(\frac{\mu_*}{\sigma_*}) / \Phi(\frac{\mu_*}{\sigma_*})$).

The EM algorithm for the stochastic frontier MHM involves the following steps:

- Step 1: Provide initial estimates of the parameter vector $\boldsymbol{\Theta}$. Set the initial value of parameters δ_1 and δ_2 equal to zero and obtain the initial value of the baseline hazard function from (7). Consistent starting values of the variances of the noise and inefficiency terms are based on method of moments estimates

$$\begin{aligned}
\hat{\sigma}_u^2 &= \left[\sqrt{2/\pi} \left(\frac{\pi}{\pi - 4} \right) \hat{m}_3 \right]^{2/3} \\
\hat{\sigma}_v^2 &= \left[\hat{m}_2 - \left(\frac{\pi - 2}{\pi} \right) \hat{\sigma}_u^2 \right]
\end{aligned} \tag{16}$$

where \hat{m}_2 and \hat{m}_3 are the estimated second and third sample moments of the OLS residuals, respectively. Estimates of σ and γ parameters are obtained through the relevant expressions provided above.

- Step 2 (E step): Compute \tilde{h}_{it} based on the current estimates and the observed data from

$$\tilde{h}_{it}^{(M)} = E \left[h_{it} | \Theta^{(M)}, \Omega_i \right] = \Pr(h_{it}^{(M)} = 1 | t_i > T_i) \begin{cases} \frac{F_e(x_i' \beta^{(M)} + \delta_1^{(M)} u_i) S_i(t; w_i, u_i)}{F_e(x_i' \beta^{(M)} + \delta_1^{(M)} u_i) S_i(t; w_i, u_i) + (1 - F_e(x_i' \beta^{(M)} + \delta_1^{(M)} u_i))} & \text{if } d_i = 0 \\ 1 & \text{otherwise} \end{cases} \quad (17)$$

- Step 3 (M step): Update the estimate of parameters by maximizing L via simulated maximum likelihood technique.
- Step 4: Iterate between steps 2 and 3 until convergence.

Continuous-time and discrete-time versions of this combined model are referred as Model III and Model IV, respectively, throughout this paper.

4 Empirical Model and Data

In this section we outline the empirical specification used in estimating the four models described above (Models I-IV). We describe the data on which our estimates are based and the step-wise forward selection procedure we employ in model building and variable selection.

4.1 Empirical Specification

Following Whalen (1991) we employ a model with a two-year timeline to estimate the probability of distress and failure and the *timing* of bank failure. In the Model I and Model III, the time to failure is measured in months²³ (1-24 months) starting from the end-year of 2007, while in the Model II and Model IV the duration times are measured in quarters as banks report their data on a quarterly basis. The covariates used in the estimation process of Model I and Model III are based on information from the fourth quarter of the 2007 Consolidated Reports of Condition and Income (Call Reports). State-specific macroeconomic variables are also derived from the Federal Reserve databases to control for state-specific effects.

We employ the cost frontier in the stochastic frontier model specification, which describes the minimum level of cost given output and input prices. The gap between the actual cost

²³Duration times measured in weeks were also considered, but not reported in this paper.

and the minimum cost is the radial measure of total (cost) inefficiency and is composed of two parts: (i) the technical inefficiency arising from excess usage of inputs and (ii) the allocative inefficiency that results from a non-optimal mix of inputs. We do not make this decomposition but rather estimate overall cost inefficiency. We adopt the intermediation approach of Sealey and Lindley (1977), according to which banks are viewed as financial intermediaries that collect deposits and other funds and transform them into loanable funds by using capital and labor. Deposits are viewed as inputs as opposed to outputs, which is assumed in the production approach.²⁴

As in Kaparakis et al., (1994) and Wheelock and Wilson (1995), we specify a multiple output-input short-run stochastic cost frontier with a quasi-fixed input. Following the standard banking literature we specify a translog functional form to describe the cost function:²⁵

$$\begin{aligned}
\log C_{it} = & \alpha_{0+} \sum_{m=1}^5 \alpha_m \log y_{mit} + \sum_{k=1}^4 \beta_k \log w_{kit} \\
& + \frac{1}{2} \sum_{m=1}^5 \sum_{j=1}^5 \alpha_{mj} \log y_{mit} \log y_{jit} + \theta_1 t + \frac{1}{2} \theta_2 t^2 \\
& + \frac{1}{2} \sum_{k=1}^4 \sum_{n=1}^4 \beta_{kn} \log w_{kit} \log w_{nit} + \eta_1 \log X_{it} + \frac{1}{2} \eta_2 (\log X_{it})^2 \\
& + \sum_{m=1}^5 \sum_{k=1}^4 \delta_{mk} \log y_{mit} \log w_{kit} + \sum_{m=1}^5 \lambda_{1x} \log y_{mit} \log X_{it} \\
& \sum_{k=1}^4 \lambda_{2x} \log w_{kit} \log X_{it} + \sum_{m=1}^5 \lambda_{mt} \log y_{mit} t + \sum_{k=1}^4 \phi_{kt} \log w_{kit} t + v_{it} + u_i
\end{aligned}$$

where C is the observed short-run variable cost of an individual bank at each time period, y_m is the value of m^{th} output, $m = 1, \dots, 5$. Outputs are real estate loans (*yreln*), commercial and industrial loans (*yciln*), installment loans (*yinln*), securities (*ysec*), and off-balance sheet items (*yobs*). The w 's represent input prices of the total interest-bearing deposits (*dep*), labor (*lab*), purchased funds (*purf*), and capital (*cap*). The quasi-fixed input (X) consists of non-interest-bearing deposits. Kaparakis et al. assume that a bank takes the level of non-interest-bearing deposits as exogenously given and since there is no market price associated with this input, the quantity of it should be included in the cost function instead of its price. We also include the time and its interaction with outputs and input prices to account for non-neutral technological change. Symmetry ($\alpha_{mj} = \alpha_{jm}$ and $\beta_{kn} = \beta_{nk}$) and linear homogeneity in input price ($\sum_{k=1}^4 \beta_k = 1$, $\sum_{k=1}^4 \beta_{kn} = \sum_{k=1}^4 \delta_{mk} = \sum_{k=1}^4 \lambda_{2x} =$

²⁴See Baltensperger (1980) for example.

²⁵The translog function provides a second-order differential approximation to an arbitrary function at a single point. It does not restrict the share of a particular input to be constant over time and across individual firms.

$\sum_{k=1}^4 \phi_{kt} = 0$) restrictions are imposed by considering capital as the numeraire and dividing the total cost and other input prices by its price.

4.2 Data

The data are from three main sources: (i) the public-use quarterly Call Reports for all U.S. commercial banks, which is collected and administrated by the Federal Reserve Bank of Chicago and the FDIC; (ii) the FDIC website, which provides information regarding failed banks and industry-level indicators; and (iii) the website of the Federal Reserve Bank of St. Louis, which provides information on regional-specific macroeconomic variables.

We drop bank observations with zero costs, zero output and input levels, as well as those with obvious measurement errors and other data inconsistencies. In addition, we exclude banks that voluntarily liquidated during the sample period and those that were chartered and started to report their data after the first quarter of 2007²⁶, which require a special treatment. The estimation sample consists of 125 banks that failed during 2008 and 2009 and 5,843 surviving banks.

More than forty bank-specific financial metrics, state-specific macroeconomic, geographical, and market structure variables are constructed from variables obtained from the above sources as potential determinants of banking distress and failure. We apply the stepwise forward selection procedure (Klein and Moeschberger, 2003) to choose the most relevant explanatory variables based on conventional statistical tests and the Akaike Information Criterion (AIC). In addition to these tests, we base our variable selection on their contribution to the overall prediction accuracy of a particular model we employ. The final set of variables pertaining to the incidence and the latency part includes proxies for the capital adequacy, asset quality, management, earnings, liquidity, and sensitivity (the so-called "CAMELS")²⁷, six market structure and geographical variables, and four state-specific variables. We use the same set of explanatory variables in both the incidence and latency parts of our models in order to capture the different effects that these have on the probability that a particular bank is troubled, as well as the probability and timing of the resolution of the bank's troubles by the FDIC. Tables 1 and 2 provide our mnemonics for the variable names, as well as their formal definitions.

The first variable in Table 1 is the tier 1 risk-based capital ratio. Banks with a high level of this ratio are considered having sufficient capital to absorb the unexpected losses occurring during the crisis and hence, have a higher chance of survival. We expect a negative sign for this variable in both the incidence and latency parts. The next variable is the ratio of nonperforming loans²⁸ to total loans, which is the primary indicator of the quality of loans made by banks and historically has been an influential factor in explaining their distress and failure. The higher this ratio, the higher the probability that the bank will enter the watch list and subsequently fail. The next five ratios also reflect banks' asset quality. We expect the ratio of allowance for loan and lease loss to average total loans to have a positive effect

²⁶These are typically referred to as the "De Novo" banks (DeYoung, 1999, 2003).

²⁷The "CAMELS" variables construction closely follows that of Lane et al. (1986) and Whalen (1991).

²⁸Nonperforming loans consist of total loans and lease financing receivables that are nonaccrual, past due 30-89 days and still accruing, and past due 90 days or more and still accruing.

on a bank’s survival. Higher ratios may signal banks to anticipate difficulties in recovering losses and thus this variable may positively impact incidence. Similarly, charge-offs on loan and lease loss recoveries provide a signal of problematic assets that increase the probability of insolvency and failure. Provision for loan and lease losses are based upon management’s evaluation of loans and leases that the reporting bank intends to hold. Such a variable can expect to decrease the probability of distress and increase the probability of survival. We can also view this as a proxy to control for one of the several ways in which different banks pursue different risk strategies (Inanoglu, et al., 2014). An often-used measure of credit risk is the gross charge-off rate (dollar gross charge-offs normalized by lending book assets). We control for risk-taking strategies in which banks may engage that differ from their role as a provider of intermediation services—the service we analyze—by including both of these risk measures as explanatory variables.

Two of the three management quality proxies that we include are constructed from the balance sheet items of the reporting banks. The first is the ratio of the full-time employees to average assets, which has an ambiguous sign in both the incidence and latency parts of our model. We conjecture, however, a negative sign on this variable as the FDIC may face constraints in seizing large banks with a large number of employees. The second is the intermediation ratio, which shows the ability of a bank to successfully transform deposits into loans and thus we expect its overall impact also to be negative. Finally, the third management quality proxy is managerial performance, which we estimate as part of our combined model. The level of banks’ earnings as measured by the operating income and returns on assets and equity are also expected to have a negative effect on both the incidence and latency parts. From liquid assets we expect cash and core deposits to have negative signs, while the direction of the effect of Jumbo time deposits is uncertain. Banks with relatively more market price sensitive liabilities and illiquid assets should be considered at a higher risk of failure *ex ante*.

5 Results and Predictive Accuracy

In Table 3, we report the results for Model I and Model II. Both models produce qualitatively similar results. The influential factors that were considered to have a strong effect on both sets of probabilities *a priori* turn out to have the correct sign and most are statistically significant in both models. Results indicate that there is a large marginal effect of the tier 1 capital ratio on the incidence probability. Other measures of earnings proxies and asset quality also have a material effect on this probability. In other words, well-capitalized banks with positive earnings and quality loans are less likely to appear on the FDIC watch list. In contrast, banks that already are on this list will increase their probability of failure if their capital ratio is insufficient, their ratio of nonperforming loans is high, and their earnings are negative and have a decreasing trend. The certificates of deposits and core deposits have the expected effect though not a statistically significant one. On the other hand, cash has a positive and significant effect. One explanation of this could be that, after controlling for profitability, banks that remain cash idle have a higher opportunity cost. It would only stand to reason for these banks to be costly and inefficient. Banks with a large number of

full-time employees are shown to have less chances to fail. Banks that successfully transform deposits into vehicles of investment are considered potentially stronger, while others with more rate sensitive liabilities appear to be less promising.

The state-specific variables have the expected economic congruences which appear to be non-significant in the incidence part of the models. We would expect these variables to significantly affect the probability of incidence of banks in the states with higher unemployment rates, lower growth in personal income, limited construction permits, and falling housing prices, all of which would give cause for increased on-site inspections. Only two of the four geographical variables have a significant effect. Banks that are Federal Reserve System (FRS) members have a higher probability of failure than those that are not. This is associated with behavior consistent with moral hazard. Such banks have felt secure as members of the FRS and hence may have assumed higher risks than they would have had they not been FRS banks. The positive result of the FRS district code indicates that the probability of insolvency and failure is higher for banks in the Atlanta (6) district than for banks in the Boston (1) district, for example. Bank size is shown to have a negative and significant effect only in the incidence part of Model II, which could be interpreted that larger banks are less likely to be placed on the watch list and subsequently fail. Finally, the older and well-established banks appear to have lower failure probabilities than their younger counterparts.

In Table 4, we present results for the continuous-time semiparametric and discrete-time MHM with the stochastic frontier specification. With few exemptions, the results are qualitatively similar to those reported in Table 3. Inefficiency has a positive effect on the incidence and failure probabilities. The effect is only significant on the latter probability and this is consistent with the view that bank performance is not the criterion for an on-site examination, but rather a factor affecting a bank's longer term viability. The distributional parameters are significant at the one-percent significance level. The descriptive statistics for the efficiency score obtained from Models III and IV as well as from the standard time-invariant random effects (RE) model for the sample of nonfailed and failed banks are summarized in Table 5. There is a small, but a statistically significant difference between the average efficiencies estimated from Models III and IV. This difference is not statistically significant for efficiencies derived from the RE model. Figure 1 depicts the distribution of inefficiencies obtained from the three models (Model III, Model IV and RE). It is worthwhile to note that the RE model reports certain surviving banks as extremely inefficient, while the most efficient banks are banks that failed. Based on these observations, we suspect that the two-step approach would yield the opposite sign on inefficiency component from what we would expect. The difference in average efficiencies from the single step estimation can be mainly attributed to the fact that distressed banks typically devote their efforts to overcome their financial difficulties and clean up their balance sheets. These impose additional costs on banks and worsen their already bad financial position.

In Figures 2 and 3 we depict the survival profile of the average bank that failed during the 2008-2009 period for all four models. It can be seen from Figure 2 that average failed banks in Model I are predicted to have a duration time of twenty two months. After controlling for inefficiencies, the time to failure drops to twenty one months. Based on the Model II results, Figure 3 demonstrates that a bank with the same characteristics as the

representative failed bank will survive up to 7 quarters, after accounting for inefficiency.

It is also interesting to look at the survival profile of the most and the least efficient banks derived from Model III and Model IV. Figure 4 displays the survival profiles obtained from Model III. The least efficient bank with an efficiency score of 0.149 percent and is predicted to fail in eight months. This bank was closed by FDIC in the end of August of 2008. On the other hand, the most efficient bank with efficiency score of 0.971 percent has a survival probability of one throughout the sample period. This is also illustrated in Figure 5, where the least efficient bank with an efficiency score of 0.154 percent is predicted to fail by fifth quarter, using the Model IV results. This bank failed in the third week of April of 2009.²⁹ The most efficient bank with an efficiency score of 0.969 percent has an estimated survival probability that exceeds 95 percent.

We next examine our results by recasting our model estimates as early warning tools that can correctly classify failed and nonfailed banks within our sample used for estimation as well as in our hold-out 2010-2011 sample. The tests are based on two types of errors, similar to those that arise in any statistical hypothesis testing. These are type I and type II errors.³⁰ A type I error is defined as the error due to classifying a failed bank as a nonfailed bank, while a type II error arises from classifying a non-failed bank as a failed bank. There is a trade-off between these two type of errors and both are important from a public policy standpoint. Models with low type I error are more desirable, since timely identification of failed banks allows the regulator to undertake any prompt corrective action to ensure that the stability and the soundness of the financial system is not compromised. On the other hand, models with high type II error unnecessary will be flagging some banks as failures while they are not, and hence could waste the regulators' time and resources. However, it is oftentimes hard to interpret the costs of a type II error since various constraints faced by the FDIC could delay the resolution of an insolvent bank. Thompson (1992) attributes this to information, administrative, legal and political constraints, among others. Whalen (1991) notes that some type II error predictions actually represent failures that occur in the near future and hence should be considered as a success of the model rather than its failure.

In Table 6, we report the in-sample predictive accuracy for the four models based on type I, type II, and overall classification error. Overall classification error is a weighted average of type I and type II errors. In what follows we set the weights at 0.5 for both errors.³¹ In our predictive accuracy analysis, each bank is characterized as a failure if its survival probability falls below a probability cutoff point, which we base on the sample average ratio of failed to nonfailed banks (0.021). The results in Table 6 indicate that the discrete-time specification yields a lower type I error than does the continuous-time specification. This is to be expected since the former incorporates multi-period observations for each bank and thus is more informative about a bank's financial health than the single-period cross-sectional observations. There is a significant drop in type I error in both specifications when the performance of a bank is added to the model as an additional factor. On the other hand

²⁹The least efficient bank is not the same in these two models. However, the most efficient bank is.

³⁰See Lane et al. 1986; Whalen, 1991; and Thompson, 1992 among others.

³¹Clearly this weighting scheme is arbitrary and alternative weighting schemes could be based on different risk preference assumptions, implicit and explicit costs of regulation, etc.

type II error is increased in the discrete-time models and it is doubled when inefficiency is included. Based on the overall classification error, Model IV performs somewhat better than Model III, but it largely outperforms the Models I and II.

Table 6 also presents the errors that judge the 2010-2011 out-of-sample classification accuracy of our models based on the estimates obtained using 2008-2009 data. The continuous-time models' errors are based on the estimated survival profiles of banks using the 2009 end-year data, while the discrete-time models' errors use the full 2010-2011 data. By comparing these errors with the 2008-2009 in-sample classification errors, we observe that there is a significant drop in type I error for all four models. This is may be due to the fact that the data used to estimate the banks' survival profiles are more informative than what was used to estimate the model parameters, which is reasonable given that the end of 2009 was considered the peek year of the 2007-2011 banking crisis. The inter-model comparison is the same as above with Model IV favored over the other models based on predictive accuracy. In addition, all four models predict the major (i.e., with total assets size over \$1 billion) and the minor bank failures equally well, by reporting very low estimated type I errors. In fact, type I error is zero for all major in- and out-of-sample bank failures.³²

In order to examine the sensitivity of the models' classification accuracy to the data period selection (high risk period versus low risk period), we also estimate the models' in-sample classification accuracy using 2010-2011 data.³³ The 2010-2011 in-sample classification errors are also summarized in Table 6. Comparing the 2010-2011 out-of-sample results to the 2010-2011 in-sample results, we observe that type I error is slightly decreased for the continuous-time models (by 0.0295 in Model I and by 0.0226 in Model III), but it is increased in the discrete-time models (by 0.041 in Model II and by 0.0176 in Model IV). More specifically, Model II fails to predict the failure of 12 out of 171 banks that failed in our 2010-2011 sample, while Model IV fails to predict the failure of 8 out of 171 failed banks during the same period.³⁴ Overall, the predictive power of our models appears to be quite robust across different estimation sub-periods within the current financial crisis. We note, however, that conditions that led to the 2007-2011 banking crisis may be substantially different from those of future banking crises. In this case, not only the model estimates, but also the variables that are used to predict banking troubles and failures can significantly differ.

5.1 Endogenous Variables and Identification

Two potential complications naturally may arise in structural models like the ones presented in this paper: the presence of endogenous variables and issues of identification of the structural model. Testing for potential endogenous control variables from our variable list and identification of the casual effect of the efficiency component is the purpose of this subsection.

³²Detailed survival profile series for each bank in our sample are available upon request.

³³The parameter estimates from the 2010-2011 estimation are available upon request.

³⁴The corresponding 2010-2011 out-of-sample predictions failed to identify only 5 of such failures.

First, we note that some of the control variables from our list of covariates may be potentially treated as endogenous in the sense that these are under a bank’s management control and potentially can be affected by the probability and timing of failure. In particular, there is the possibility that a bank that is placed on the FDIC’s watch list may erroneously report (underestimate or overestimate) the amount of these variables in its Call Reports. Such variables may include the provision for loan and lease losses³⁵, which involves subjective assessment by a bank’s management, and the number of the full-time employees, which is subject to substantial variation during distressed times. Other variables, such as allowance for loan and lease loss, charge-offs on loans and leases and recoveries on allowance for loan and lease losses also can be treated as endogenous. However, we note that these are subject to stringent scrutiny by regulators and auditors who can recognize and measure the effectiveness and appropriateness of management’s methodology for collectively and individually assessing these accounts in accordance with internationally accepted reporting standards. We, therefore, treat these variables as exogenous in our models and do not further test for their exogeneity.

Below we do test for the endogeneity of the provision for loan and lease losses and the number of the full-time employees. We use a nonparametric test based on Abrevaya, Hausman and Khan (2010). The test is carried out using the following steps:

- Step 1: Identify, select and validate instrumental variables for the potentially endogenous variables;
- Step 2: Project the potentially endogenous variables onto the column space of the instrumental and exogenous variables and obtain their fitted values;
- Step 3: Estimate the model³⁶ separately by using the potentially endogenous variables and instrumented endogenous variables and obtain the survival profiles under both cases (label these as S_{end} and S_{iv} , respectively)
- Step 4: Use Kendall’s tau rank correlation statistic to test for association/dependence of S_{end} and S_{iv} (i.e., test for the null that S_{end} and S_{iv} are not associated/dependent)
- Step 5: Reject the null hypothesis of endogeneity if the p-value of Kendall’s tau statistic is below the desired confidence level.

For the provision for loan and lease losses/average loans and leases variable, the selected instruments are (i) one period lagged values of the provision for loan and lease losses/average loans and leases; (ii) one period lagged values of the nonperforming loans/total loans; (iii) one period lagged values of the allowance for loan and lease loss/average loans and leases; and (iv) one period lagged values of the recoveries on allowance for loan and lease

³⁵The provision for loan and lease loss is the amount required to establish a balance in the allowance for credit losses account, which management considers adequate to absorb all credit related losses in its loan portfolio.

³⁶Note that for testing purposes only the time-varying model combined with efficiencies (i.e., Model IV) is used.

losses/average loans and leases. The estimated Kendall's tau statistic is 0.9293 (with p-value = 0) and thus we reject the null hypothesis that the provision for loan and lease losses variable is endogenous in our estimation sample. Similarly, for the number of full-time equivalent employees/average assets variable, the selected instruments are (i) current period overhead expense; (ii) one period lagged values of the overhead expense; (iii) current period ratio of non-interest expense/total assets; (iv) one period lagged values of the ratio of non-interest expense/total assets; (v) current period ratio of non-interest expense/interest expense; and (vi) one period lagged values of the ratio of non-interest expense/interest expense. The estimated Kendall's tau statistic is 0.9723 (with p-value = 0) and we thus reject the null hypothesis that the number of the full-time equivalent employees variable is endogenous in our estimation sample. Joint testing yields Kendall's tau statistic of 0.9115 (with p-value = 0), thus leading to the same conclusion that both of these variables are not endogenous in our sample.

To corroborate the testing results above, we also test for the endogeneity of the provision for loan and lease losses and the number of the full-time employees by considering only the incidence part of the model. The rationale for using this alternative testing approach is that one might consider that these variables would be affected primarily by the incidence probability, as a bank's management could potentially manipulate these accounts to avoid being placed on the FDIC's watch list in the first place. The testing results are based on the Wald statistic on the hypothesis is exogeneity of the potential endogenous variables.³⁷ The Wald statistic is 1.74 (with p-value = 0.1866) for the provision for loan and lease losses/average loans and leases and 3.06 (with p-value = 0.0804) for the number of the full-time employees/average assets. These estimated test statistics are not significant at the 5% confidence level and generally corroborate the findings using the alternative null hypothesis.

The identification of the casual effect of the efficiency term, on the other hand, is performed by testing for the over-identifying restrictions using the testing approach outlined above. Due to the fact that the efficiency term is latent (unobserved) in our models, we use the efficiency scores obtained from the random effects (RE) model as a proxy for the combined model's efficiencies. We identify the one period lagged values of the return on assets, the one period lagged values of the return on equity, the one period lagged value of the intermediation ratio (total loans/total deposits), the ratio of non-interest expense/interest expense, and the one period lagged values of the ratio of non-interest expense/interest expense as instrumental variables for the estimated efficiency scores. The resulting Kendall's tau statistic is estimated as 0.7970 (with p-value = 0); thus, rejecting the null hypothesis that the casual effects are not identified in our estimation sample.

6 Concluding Remarks

Massive banking failures during the financial turmoil of the Great Recession has resulted in enormous financial losses and costs to the U.S. economy, not only in terms of bailouts by regulatory authorities in their attempt to restore liquidity and stabilize the financial sector,

³⁷This testing is carried out by using STATA's *ivprobit* command.

but also in terms of lost jobs in banking and other sectors of economy, failed businesses, and ultimately slow growth of the economy as a whole. The design of early warning models that accurately predict the failures and their timing is of crucial importance in order to ensure the safety and the soundness of the financial system. Early warning models that can be used as off-site examination tools are useful for at least three reasons. They can help direct and efficiently allocate the limited resources and time for on-site examination so that banks in immediate help are examined first. Early warning models are less costly than on-site visits made by supervisors to institutions considered at risk and can be performed with high frequency to examine the financial condition of the same bank. Finally, early warning models can predict failures at a reasonable length of time prior to the marked deterioration of bank's condition and allow supervisors to undertake any prompt corrective action that will have minimal cost to the taxpayer.

In this paper we have considered early warning models that attempt to explain recent failures in the U.S. commercial banking sector. We employed a duration analysis model combined with a static logit model to determine troubled banks which subsequently fail or survive. Both continuous and discrete time versions of the mixed model were specified and estimated. These effectively translated the bank-specific characteristics, state-related macroeconomic variables, and geographical and market structure variables into measures of risk. Capital adequacy and nonperforming loans were found to play a pivotal role in determining and closing insolvent institutions. State-specific variables appeared to significantly affect the probability of failure but not insolvency. The discrete-time model outperformed the continuous-time model as it is able to incorporate time-varying covariates, which contain more and richer information. We also found that managerial efficiency does not significantly affect the probability of a bank being troubled but plays an important role in their longer term survival. Inclusion of the efficiency measure led to improved prediction in both models.

7 Appendix

In this appendix we show the derivation of the sample likelihood function given in expression (3). For this purpose we first note that at time t , bank i can fall into four mutually exclusive states of nature:

$$States = \begin{cases} h_i = 1, d_i = 1 \text{ (Problem \& Failed)} & \text{with prob. } F_e(x'_i \boldsymbol{\beta}) \lambda_i^p(t; w_i) S_i^p(t; w_i) \\ h_i = 0, d_i = 1 \text{ (Sound \& Failed)} & \text{with prob. } [1 - F_e(x'_i \boldsymbol{\beta})] \lambda_i^s(t; w_i) S_i^s(t; w_i) \\ h_i = 1, d_i = 0 \text{ (Problem \& Censored)} & \text{with prob. } F_e(x'_i \boldsymbol{\beta}) S_i^p(t; w_i) \\ h_i = 0, d_i = 0 \text{ (Sound \& Censored)} & \text{with prob. } [1 - F_e(x'_i \boldsymbol{\beta})] S_i^s(t; w_i) \end{cases}$$

Then

$$\begin{aligned} L(\boldsymbol{\theta}; x, w, d) &= \prod_{i=1}^n L_i(\boldsymbol{\theta}; x, w, d) \\ &= \prod_{i=1}^n \left\{ [F_e(x'_i \boldsymbol{\beta}) \lambda_i^p(t; w_i) S_i^p(t; w_i)]^{h_i} \left([1 - F_e(x'_i \boldsymbol{\beta})] \lambda_i^s(t; w_i) S_i^s(t; w_i)^{1-h_i} \right) \right\}^{d_i} \\ &\quad \times \left\{ [F_e(x'_i \boldsymbol{\beta}) S_i^p(t; w_i)]^{h_i} [1 - F_e(x'_i \boldsymbol{\beta})] S_i^s(t; w_i)^{1-h_i} \right\}^{1-d_i} \\ &= \prod_{i=1}^n F_e(x'_i \boldsymbol{\beta})^{h_i} [1 - F_e(x'_i \boldsymbol{\beta})]^{(1-h_i)} [\lambda_i^p(t; w_i)]^{d_i h_i} \\ &\quad \times [\lambda_i^s(t; w_i)]^{d_i(1-h_i)} [S_i^p(t; w_i)]^{h_i} [S_i^s(t; w_i)]^{1-h_i} \end{aligned}$$

By assumption, $\lambda_i^s(t; w_i) = 0$, if and only if, $h_i = 0$ and $d_i = 0$ (i.e., a bank is healthy and is not observed failing). Similarly $S_i^s(t; w_i) = 1$ if and only if $h_i = 0$ (i.e., a bank is healthy). The final sample likelihood function is then given by

$$L(\boldsymbol{\theta}; x, w, d) = \prod_{i=1}^n F_e(x'_i \boldsymbol{\beta})^{h_i} [1 - F_e(x'_i \boldsymbol{\beta})]^{(1-h_i)} [\lambda_i^p(t; w_i)]^{d_i h_i} [S_i^p(t; w_i)]^{h_i}$$

which implies that the completely healthy banks contribute to the likelihood function only through their probability being troubled.

Table 1: CAMELS proxy Financial Ratios

| Capital Adequacy (C) | |
|------------------------|---|
| tier1 | Tier 1 (core) capital/risk-weighted assets |
| Asset Quality (A) | |
| rnpl | Nonperforming loans/total loans |
| all | Allowance for loan and lease loss/average loans and leases |
| reln | Commercial real estate loans/total loans |
| coffs | Charge-offs on loans and leases/average loans and leases |
| lrec | Recoveries on allowance for loan and lease losses/average loans and leases |
| llp | Provision for loan and lease losses /average loans and leases |
| Managerial Quality (M) | |
| fte | (Number of fulltime equivalent employees/average assets)*1000 |
| imr | Total loans/total deposits |
| u | Random Effects inefficiency score |
| Earnings (E) | |
| oi | Total operating income/average assets |
| roa | Net income (loss)/average assets |
| roe | Net income (loss)/total equity |
| Liquidity (L) | |
| cash | Noninterest-bearing balances, currency, and coin/average assets |
| cd | Total time deposits of USD 100,000 or more/total assets |
| coredep | Core deposits/total assets |
| Sensitivity (S) | |
| sens | Difference in interest rate sensitive assets and liabilities repricing within one year/total assets |

Table 2: Geographical, Market Structure, and State-Specific Macroeconomic Variables

| Geographical and Market Structure Variables | |
|---|---|
| chtype | Charter type (1 if state chartered, 0 otherwise) |
| frsmb | FRS membership indicator (1 if Federal Reserve member, 0 otherwise) |
| ibf | International banking facility (1 if bank operates an international based facility, 0 otherwise) |
| frsdistrcode | FRS district code: (Boston(1), New York (2), Philadelphia (3), Cleveland (4), Richmond (5), Atlanta (6), Chicago (7), St. Louis (8), Minneapolis (9), Kansas City (10), Dallas (11), San Francisco (12), Washington, D.C. (0-reference district)) |
| lgta | log of total assets |
| age | Age (measured in months or quarters) |
| State-Specific Macroeconomic variables | |
| ur | Unemployment rate |
| chpi | Percentage change in personal income |
| chphi | Percentage change in house price index |
| chnphu | Change in new private housing units authorized by building permits |

Table 3: Parameter Estimates obtained under Model I (SPMHM) and Model II (DTMHM)

| Variable | Model I | | Model II | |
|--------------|---------------------|---------------------|---------------------|---------------------|
| | Latency | Incidence | Latency | Incidence |
| Intercept | | -2.5989 (2.8512) | | 4.9130 (2.9266) |
| lgta | 0.0797 (0.0885) | 0.0607 (0.1103) | 0.0531 (0.0875) | -0.3320*** (0.1056) |
| age | -0.0004* (0.0002) | 0.0004 (0.0003) | -0.0003 (0.0002) | 0.0001 (0.0003) |
| tier 1 | -48.417*** (3.0567) | -86.791*** (5.3156) | -47.060*** (3.0856) | -88.728*** (5.3516) |
| alll | -9.5829** (4.7047) | 16.473** (7.9759) | -8.8615* (4.6962) | 8.8671 (7.7594) |
| reln | 4.4321*** (1.1801) | 2.0116 (1.2748) | 3.7811*** (1.1731) | 3.9762*** (1.2796) |
| rnpl | 7.2555*** (1.3348) | 6.3838*** (2.1574) | 6.1802*** (1.3433) | 9.6447*** (2.1510) |
| roa | -6.1672 (5.1795) | -11.248** (5.5416) | -7.2727 (5.0983) | -8.8145 (6.1201) |
| roe | 0.0003 (0.0003) | 0.0002 (0.0013) | 0.0003 (0.0004) | 0.0003 (0.0017) |
| cd | 1.0098 (0.8644) | 1.6651 (1.0274) | 1.2499 (0.8425) | 0.8245 (1.0003) |
| coredep | -2.7654 (1.7546) | -1.2140 (2.0839) | -2.5466 (1.7496) | -3.1272 (2.0927) |
| coffs | 0.2351*** (0.0804) | 0.3168*** (0.1183) | 0.2319*** (0.0848) | 0.2703** (0.1243) |
| lrec | 38.162** (18.672) | 14.463 (56.448) | 35.681* (21.726) | 37.945 (42.003) |
| llp | -10.427** (4.9577) | -15.501** (6.5158) | -11.688** (4.8628) | -13.155** (6.6190) |
| fte | -0.8228 (1.0468) | -3.0004** (1.4396) | -0.8329 (1.0512) | -3.1287** (1.4021) |
| imr | -4.2141*** (1.0634) | -1.7238 (1.2254) | -3.7792*** (1.0603) | -4.4020*** (1.2016) |
| sens | 2.3255*** (0.8386) | 2.5869** (1.0320) | 2.0025** (0.8403) | 5.6444*** (1.0042) |
| cash | 6.7983*** (2.0542) | 6.7497** (3.2628) | 6.9211*** (2.0472) | 4.5465 (3.6333) |
| oi | -3.9670 (4.4353) | -6.1756 (6.6955) | -3.1670 (4.3948) | -4.9651 (6.6585) |
| ur | 0.1198*** (0.0379) | 0.0196 (0.0490) | 0.0655* (0.0390) | 0.0548 (0.0482) |
| chpi | -15.091* (8.1323) | -10.555 (9.7490) | -20.313** (8.0823) | -10.645 (10.017) |
| chhpi | -8.1375* (4.9453) | -3.1678 (5.8411) | -9.8817** (4.8428) | -5.4824 (5.8215) |
| chnphu | -0.6570*** (0.2490) | 0.0006 (0.0523) | -0.5246** (0.2417) | 0.0047 (0.0581) |
| chtype | -0.2151 (0.5058) | 0.4441 (0.6871) | 0.0223 (0.5051) | -0.7143 (0.5943) |
| frsmb | 0.4707*** (0.1797) | 0.4018* (0.2352) | 0.4617*** (0.1808) | 0.3466 (0.2363) |
| ibf | 1.1171 (0.7589) | 1.4405 (0.8883) | 1.2816* (0.7592) | -2.5959*** (0.7825) |
| frsdistrcode | 0.2465*** (0.0329) | 0.2615*** (0.0430) | 0.2295*** (0.0325) | 0.2457*** (0.0427) |
| LogL | 1,763.87 | | 1,714.92 | |
| N | 5,968 | | 38,571 | |

p* < 0.1, p** < 0.05, p*** < 0.01 (Robust standard errors in parentheses)

Table 4: Parameter Estimates obtained under Model III (SPMHM+SF) and Model IV (DTMHM+SF)

| Variable | Model III | | Model IV | |
|--------------|-----------------------|---------------------|---------------------|---------------------|
| | Latency | Incidence | Latency | Incidence |
| Intercept | | -2.6934 (2.8039) | | 4.7694* (2.9201) |
| lgta | -0.0408 (0.0935) - | -0.0087 (0.1243) | -0.0742 (0.0958) | 0.3466*** (0.1117) |
| age | -0.0004* (0.0002) | 0.0004 (0.0003) | -0.0004* (0.0002) | 0.0001 (0.0003) |
| tier 1 | -48.647*** (3.0631) | -86.280*** (5.3068) | -48.452*** (3.0678) | -88.684*** (5.3391) |
| alll | -8.5073* (4.6488) | 17.003** (7.9738) | -8.8587* (4.6066) | 8.8881 (7.7741) |
| reln | 4.6588*** (1.1259) | 2.1044* (1.2631) | 4.4871*** (1.0878) | 3.9288*** (1.2805) |
| rnpl | 6.9014*** (1.3206) | 6.0653*** (2.1661) | 6.7347*** (1.3210) | 9.5835*** (2.1554) |
| roa | -6.4175 (5.0868) | -11.451*** (5.5328) | -6.1672 (5.1423) | -8.8129 (6.1180) |
| roe | 0.0002 (0.0004) | 0.0001 (0.0013) | 0.0002 (0.0003) | 0.0002 (0.0017) |
| cd | 0.8641 (0.8608) | 1.5840 (1.0277) | 0.7565 (0.8579) | 0.7329 (1.0244) |
| coredep | -2.3913 (1.6224) | -1.0244 (2.0568) | -1.5432 (1.6367) | -2.9196 (2.1285) |
| coffs | 0.2447*** (0.0801) | 0.3232*** (0.1172) | 0.2516*** (0.0798) | 0.2720** (0.1243) |
| lrec | 37.309** (19.148) | 14.661 (56.167) | 37.219** (19.011) | 38.569 (41.568) |
| llp | -11.175** (4.9577) | -15.784** (6.5199) | -11.654** (4.7671) | -13.211** (6.6345) |
| fte | -2.1781** (1.0195) | -3.8780** (1.5932) | -2.8670*** (1.0298) | -3.3559** (1.5165) |
| imr | -3.7660*** (0.9728) - | -1.4553 (1.2128) | -3.2640*** (0.9832) | 4.2466*** (1.2601) |
| sens | 2.2143*** (0.8294) | 2.5264** (1.0309) | 2.0894** (0.8282) | 5.6036*** (1.0065) |
| cash | 7.4166*** (2.0368) | 7.1461*** (3.2558) | 7.6605*** (2.0375) | 4.6012 (3.6396) |
| oi | -3.9483 (4.3968) | -6.4722 (6.6946) | -4.1980 (4.3825) | -5.0126 (6.6601) |
| ur | 0.1210*** (0.0377) | 0.0234 (0.0491) | 0.1208*** (0.0378) | 0.0555 (0.0487) |
| chpi | -15.567** (8.1081) | -9.7061 (9.7811) | -15.551* (8.0642) | -10.639 (10.021) |
| chhpi | -8.1886* (4.9593) | -3.2387 (5.8526) | -8.1802** (4.9731) | -5.4864 (5.8139) |
| chnphu | -0.6300*** (0.2471) | -0.0001 (0.0531) | -0.6171*** (0.2456) | 0.0046 (0.0578) |
| chtype | -0.1496 (0.5045) | 0.4875 (0.6876) | -0.1293 (0.5028) | -0.7224 (0.5953) |
| frsmb | 0.4960** (0.1801) | 0.3994* (0.2349) | 0.4977*** (0.1801) | 0.3487 (0.2512) |
| ibf | 1.1325 (0.7574) | 1.4718* (0.8945) | 1.1295* (0.7571) | -2.5923*** (0.7811) |
| frsdistrcode | 0.2612*** (0.0332) | 0.2725*** (0.0445) | 0.2663*** (0.0334) | 0.2469*** (0.0425) |
| δ_1 | | 0.2062 (0.1577) | | 0.0343 (0.0828) |
| δ_2 | 0.3058*** (0.0468) | | 0.4137*** (0.0750) | |
| σ | | 0.0552*** (0.0011) | | 0.0548*** (0.0011) |
| γ | | 0.5173*** (0.0017) | | 0.5278*** (0.0015) |
| LogL | | 67,701 | | 66,360 |
| N | | 5,968 | | 38,571 |

p* < 0.1, p** < 0.05, p*** < 0.01 (Robust standard errors in parentheses)

Table 5: Cost Efficiencies Results

| | Mean | Standard Deviation | Minimum | Maximum |
|-----------------|--------|--------------------|---------|---------|
| NonFailed Banks | | | | |
| Model III | 0.6817 | 0.0691 | 0.3167 | 0.9705 |
| Model IV | 0.7295 | 0.1630 | 0.1992 | 0.9688 |
| Random Effects | 0.6466 | 0.0662 | 0.4636 | 0.9650 |
| Failed Banks | | | | |
| Model III | 0.6721 | 0.1022 | 0.1499 | 0.8722 |
| Model IV | 0.6804 | 0.0824 | 0.1539 | 0.8488 |
| Random Effects | 0.6408 | 0.0798 | 0.3845 | 0.8626 |

The top and bottom 5% of inefficiencies scores are trimmed to remove the effects of outliers

Table 6: Predictive Accuracy Results

| | Model I | Model II | Model III | Model IV |
|--|---------|----------|-----------|----------|
| 2008-2009 in-sample classification | | | | |
| Type I error | 0.3840 | 0.2882 | 0.1123 | 0.0644 |
| Type II error | 0.0047 | 0.0051 | 0.0231 | 0.0476 |
| Overall classification error | 0.1937 | 0.1465 | 0.0581 | 0.0573 |
| 2010-2011 out-of-sample classification | | | | |
| Type I error | 0.2283 | 0.0292 | 0.1630 | 0.0292 |
| Type II error | 0.0049 | 0.0012 | 0.0062 | 0.0012 |
| Overall classification error | 0.1157 | 0.0152 | 0.0840 | 0.0152 |
| 2010-2011 in-sample classification | | | | |
| Type I error | 0.1988 | 0.0702 | 0.1404 | 0.0468 |
| Type II error | 0.0025 | 0.0012 | 0.0025 | 0.0012 |
| Overall classification error | 0.1007 | 0.0357 | 0.0715 | 0.0240 |

Overall classification error is a simple average of type I and type II errors

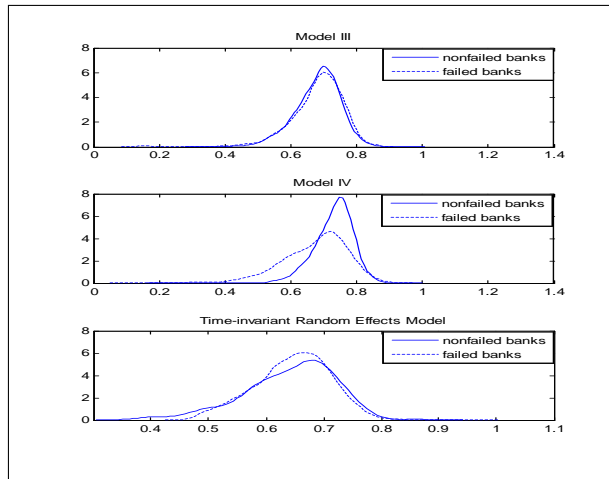


Figure 1: Cost Efficiency Distributions

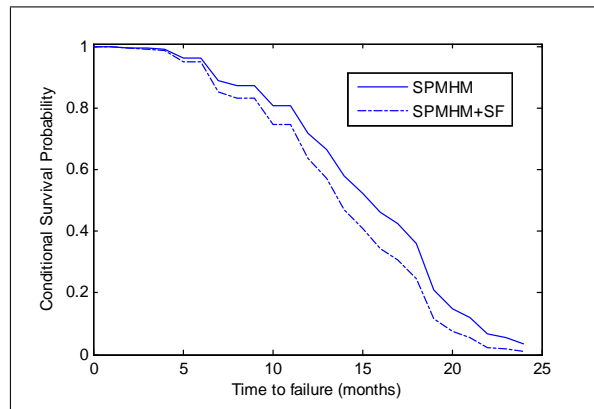


Figure 2: SPMHM: Failed Banks' Average Survival Profile

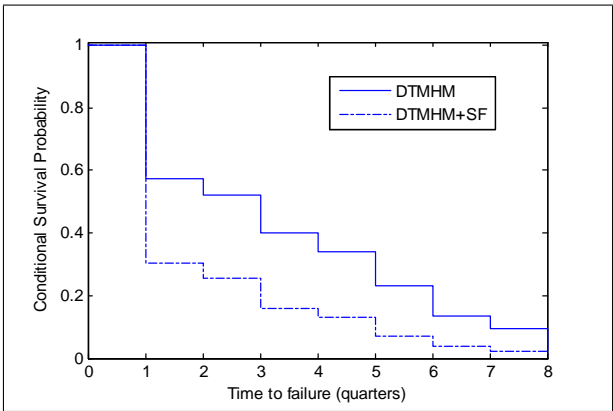


Figure 3: DTMHM: Failed Banks' Average Survival Profile

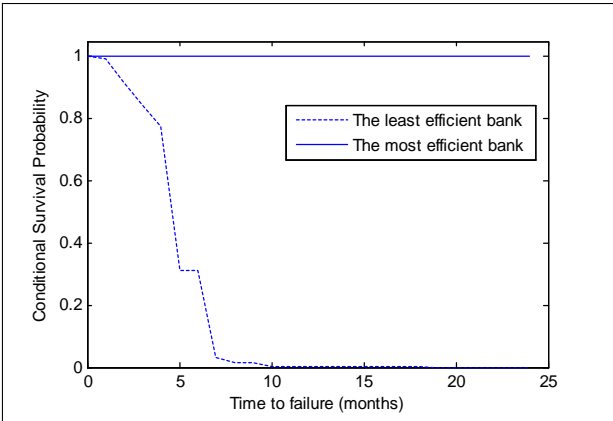


Figure 4: Model III: The Most and the Least Efficient Bank's Survival Profile

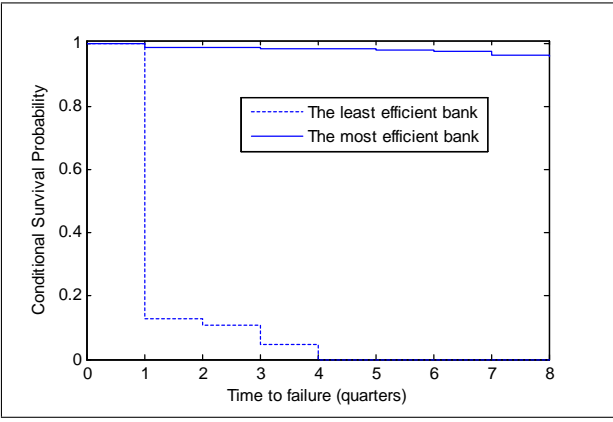


Figure 5: Model IV: The Most and the Least Efficient Bank's Survival Profile

References

- [1] Abrevaya, J, Hausman, J.A., and Khan, S., 2010, Testing for Causal Effects in a Generalized Regression Model With Endogenous Regressors. *Econometrica* 78 (6), 2043–2061.
- [2] Aigner, D.J., Lovell , C.A.K., and Schmidt, P., 1977, Formulation and estimation of stochastic frontier models. *Journal of Econometrics* 6, 21–37.
- [3] Altman, E. I., 1968, Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy. *Journal of Finance* 23, 589–609.
- [4] Baltensperger, E., 1980, Alternative approaches to the theory of the banking firm. *Journal of Monetary Economics* 6, 1–37.
- [5] Barr, R.S., and Siems, T.F., 1994, Predicting bank failure using DEA to quantify management quality. *Financial Industry Studies Working Paper* 94–1, Federal Reserve Bank of Dallas.
- [6] Battese, G.E., and Cora, G.S., 1977, Estimation of a production frontier model: with application to the pastoral zone of eastern Australia. *Australian Journal of Agricultural Economics* 21, 169–179.
- [7] Bover, O., Arellano, M., and Bentolila, S., 2002, Unemployment duration, benefit duration and the business cycle. *Economic Journal* 112, 223–265.
- [8] Charnes, A., Cooper, W.W., and Rhodes, E.L., 1978, Measuring the efficiency of decision making units. *European Journal of Operational Research* 2, 429–444.
- [9] Cole, R.A., and Gunther, J.W., 1995, Separating the likelihood and timing of bank failure. *Journal of Banking and Finance* 19, 1073–1089.
- [10] Cole, R.A., and Gunther, J.W., 1998, Predicting bank failures: A comparison of on- and off-site monitoring systems. *Journal of Financial Services Research* 13, 103–117.
- [11] Cole, R.A., and Wu, Q., 2009, Predicting bank failures using a simple dynamic hazard model. FDIC Working Paper, Washington, D.C.
- [12] Cole, R.A., and Wu, Q., 2011, Is hazard or probit more accurate in predicting financial distress? Evidence from U.S. bank failures. MPRA Paper No. 29182, Munich, Germany.
- [13] Cornwell, C., Schmidt, P., Sickles, R.C., 1990, Production frontiers with cross-sectional and time series variation in efficiency levels. *Journal of Econometrics* 46, 185–200.
- [14] Cox, D.R., 1972, Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society, Series B* 34, 187–220.
- [15] Deakin, E., 1972, A discriminant analysis of predictors of business failure. *Journal of Accounting Research*, Spring, 167–179.

- [16] Debreu, G., 1951, The coefficient of resource utilisation. *Econometrica* 19, 273–292.
- [17] Dempster, A.P., Laird, N.M., and Rubin, D.B., 1977, Maximum likelihood from incomplete data via the em algorithm (with discussion). *Journal of the Royal Statistical Society, Series B* 39, 1–38.
- [18] DeYoung, R., 1999, Birth, growth, and life or death of newly chartered banks. *Economic Perspectives* 23, 18–35.
- [19] DeYoung, R., 2003, The Failure of new entrants in commercial banking markets: a split-population duration analysis. *Review of Financial Economics* 12, 7–33.
- [20] Farewell, V.T., 1977, A model for a binary variable with time-censored observations. *Biometrika* 64, 43–46.
- [21] Farewell, V.T., 1982, The use of mixture models for the analysis of survival data with long-term survivor. *Biometrics* 38, 1041–1046.
- [22] Farrell, M., 1957, The measurement of productive efficiency. *Journal of the Royal Statistical Society A, General*, 120, 253–281.
- [23] Greene, W.H., 2008, The econometric approach to efficiency analysis, chapter 2. In *The Measurement of Productive Efficiency: Techniques and Applications*, . H. O. Fried, C. A. K. Lovell and S.S. Schmidt (eds.), New York: Oxford University Press.
- [24] Halling, M. and Hayden, E., 2007, Bank failure prediction: a two-step survival time approach. *Proceedings of the International Statistical Institute’s 56th Session*.
- [25] Hausman, J.A, and Taylor, W. E., 1981, Panel data and unobservable individual effects. *Econometrica* 49, 1377–1398.
- [26] Inanoglu, H., Jacobs, M., Liu, R., and Sickles, R.C., 2014, Analyzing Bank Efficiency: Are “too-big-to-fail” Banks Efficient?, Rice University, mimeo.
- [27] Jondrow, J., Lovell, C.A.K., Materov, I.S., Schmidt, P., 1982, On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, 19, 233–238.
- [28] Kalbfleisch, J. D., and Prentice, R.L., 2002, *The Statistical Analysis of Failure Time Data*. 2nd ed. New York: Wiley.
- [29] Kaparakis, E.I., Miller, S.M., and Noulas, A.G., 1994, Short-run cost inefficiency of commercial banks: a flexible stochastic frontier approach. *Journal of Money, Credit and Banking* 26, 875–893.
- [30] Kasa, K., and Spiegel, M.M., 2008, The role of relative performance in bank closure decisions. *Economic Review*, Federal Reserve Bank of San Francisco.
- [31] Klein, J.P., and Moeschberger, M.L., 2003, *Survival Analysis. Techniques for Censored and Truncated Data*, 2nd Edition, New York: Springer.

- [32] Kuk, A.Y C., and Chen, C.H., 1992, A mixture model combining logistic regression with proportional hazards regression. *Biometrika*, 79, 531–541.
- [33] Kumbhakar, S., and Lovell, C.A.K., 2000, *Stochastic Frontier Analysis*. Cambridge, MA: Cambridge University Press.
- [34] Lancaster, T., 1990, *The Econometric Analysis of Transition Data*. Cambridge, MA: Cambridge University Press.
- [35] Lane, W., Looney, S., and Wansley, J., 1986, An application of the Cox proportional hazards model to bank failure. *Journal of Banking and Finance* 10, 511–531.
- [36] Lu, W., 2008, Maximum likelihood estimation in the proportional hazards cure model. *Annals of the Institute for Statistics and Mathematics* 60: 545–574.
- [37] Martin, D., 1977, Early warning of bank failure: A logit regression approach. *Journal of Banking and Finance* 1, 249–276.
- [38] McLachlan, G., and Krishnan, T., 1996, *The EM Algorithm and Extensions*. New York: John Wiley & Sons.
- [39] Meeusen, W., and van den Broeck, J., 1977, Efficiency estimation from Cobb–Douglas production functions with composed error. *International Economic Review* 18, 435–444.
- [40] Meyer, B., 1990, Unemployment insurance and unemployment spells. *Econometrica* 58, 757–782.
- [41] Meyer, P.A., and Pifer, H.W., 1970, Prediction of bank failures. *Journal of Finance* 25, 853–868.
- [42] Pitt, M., and Lee, L., 1981, The measurement and sources of technical inefficiency in the Indonesian weaving industry. *Journal of Development Economics*, 9, 43–64.
- [43] Sealey, S., and Lindley, J.T., 1977, Inputs, outputs, and a theory of production and cost at depository financial institutions. *Journal of Finance* 32, 1251–1266.
- [44] Shumway, T., 2001, Forecasting bankruptcy more accurately: a simple hazard model. *Journal of Business* 74, 101–124.
- [45] Sickles, R. C., and Taubman, P., 1986, An analysis of the health and retirement status of the elderly. *Econometrica* 54, 1339–1356.
- [46] Sy, L., and Taylor, J., 2000, Estimation in a Cox proportional hazards cure model. *Biometrics*, 56, 227–236.
- [47] Thompson, J.B., 1992, Modeling the regulator’s closure option: a two-step logit regression approach. *Journal of Financial Services Research* 6, 5–23.

- [48] Topaloglu, Z., and Yildirim, Y., 2009, Bankruptcy Prediction. Working Paper, Graduate Center, CUNY, Department of Economics.
- [49] Torna, G., 2010, Understanding Commercial Bank Failures in the Modern Banking Era. Available at: <http://www.fma.org/NY/Papers/ModernBanking-GTORNA.pdf>
- [50] Tsionas, E.G., and Papadogonas, T.A., 2006, Firm exit and technical inefficiency. *Empirical Economics* 31, 535–548.
- [51] Whalen, G., 1991, A proportional hazards model of bank failure: An examination of its usefulness as an early warning model tool. *Economic Review*, Federal Reserve Bank of Cleveland, 21–31.
- [52] Wheelock, D. and Wilson, P., 1995, Explaining bank failures: deposit insurance, regulation, and efficiency. *Review of Economics and Statistics* 77, 689–700.
- [53] Wheelock, D and Wilson, P., 2000, Why do banks disappear? the determinants of U.S. bank failures and acquisitions. *Review of Economics and Statistics* 82, 127–138.
- [54] Yildirim, Y., 2008, Estimating default probabilities of cmbs with clustering and heavy censoring. *The Journal of Real Estate Finance and Economics* 37, 93–111.