Agricultural Banking and Bank Failures of the Late 2000s Financial Crisis: A Survival Analysis Using Cox Proportional Hazard Model

Xiaofei Li
University of Georgia, Department of Ag & Applied Economics
306 Conner Hall, Athens GA 30602
Email: zxmf1128@uga.edu

Cesar L. Escalante
University of Georgia, Department of Ag & Applied Economics
315 Conner Hall, Athens GA 30602
Email:cescalan@uga.edu

James E. Epperson
University of Georgia, Department of Ag & Applied Economics
315 Conner Hall, Athens GA 30602
Email: epperson@uga.edu

Selected Paper prepared for presentation at the Southern Agricultural Economics Association (SAEA) Annual Meeting, Dallas, Texas, 1 - 4 February 2014

Abstract

This study employs a semi-parametric Cox proportional hazard model to study the relationship between survival time and bank-specific determinants of failure of commercial and agricultural banks during the recent recessionary period. Results indicate that non-performing consumer and commercial loans have seriously impaired banks' financial health and survival.

Key Words: Agricultural bank, Bank failure, Cox proportional hazard model, Survival analysis

Introduction

As declared by the National Bureau of Economic Research, the late 2000s economic recession that affected both the U.S. and global economies is considered as the longest economic downturn since the 1930s Great Recession (NBER, 2010). This recession, characterized by high unemployment, declining real estate values, bankruptcies and foreclosures, has affected the banking industry so hard that a surge of bank failures occurred in the United States. Following 25 bank closures in 2008, a total of 140 banks shut down in 2009. Although the economy has already shown signs of recovery since the recession was said to have ended on June 2009, the rate of bank closure even increased in 2010 with 157 bank failures, the highest level since the savings-and-loan crisis in 1992. This trend was followed by 92 more bank failures in 2011 and 51 in 2012. By October 2013, a total of 488 banks failed during the last five years.

Investments in residential mortgage-backed securities (RMBS) have been singled out as having triggered this latest financial crisis. A dramatic increase in delinquencies in subprime residential loan accommodations due to the housing boom-and-bust in 2006 has caused the default by hundreds of thousands of borrowers within a short period of time and resulted in a numbers of banks, particularly those highly involved in the RMBS market, closing down or being taken over due to their insufficient capital and incapability to survive the ensuing financial distress.

A crisis in all or a part of the banking sector may result in a decline in shareholders' equity value, the loss of depositors' savings, and insufficient funding for borrowers. These would translate to increasing costs on the economy as a whole or parts within it (Hoggarth et al. 2002). In this regard, it is important to probe more deeply and understand the causes of bank failures, which should provide insights on more effective solutions to the current economic crises or cautionary policies that will prevent its duplication in the future. Thus, the detection of early warning signals of bank's tendency to fail can help to modify banks' operating decisions and strategies, and help a bank avoid failure in the future.

Meanwhile, compared to regular commercial banks, agricultural banks[1] usually have more liquidity concerns. Thus, they are unable to diversify their clientele to include other non-agricultural business clientele due to funding constraints. The specialized nature of their lending operations and the large variability of the agricultural products' prices usually result in greater risks and uncertainty. However, during the financial crisis, agricultural banks have been generally in stronger financial health since most of the agricultural-related financial institutions did not participate aggressively in the commercial real estate industry and agricultural banks did not invest in the structured securities that have lost substantial market value. Thus, apart from the detection of early-warning signals for commercial banks, it is also important to identify the factors that enhanced the survival ability of agricultural banks.

Among the large number of early warning studies that have already been published, most have employed probit/logit technique to construct the models (Cole and Gunther, 1998, Hanweck, 1977, Martin, 1977, Pantalone and Platt, 1987, Thomson, 1991). The adoption of a binary model design allows us to divide the banks into two classes: failure and non-failure, and generate a bank's probability of falling into one group or the other with a given set of bank characteristic variables. The use of duration models to explain and predict bank failure is relatively more recent approach compared to the basis probit/logit technique. Cole and Gunther (1995), Lane et al

[1] FDIC defines Agricultural banks as "Banks whose agricultural production loans plus real estate loans secured by farmland exceed 25 percent of total loans and leases".

(1986), and Weelock and Wilson (1995) have used different duration techniques in their studies related to bank failure. The duration model has been preferred over probit/logit model given its capability to generate not just estimates of the probability of bank failure but also estimates of the probable time to failure.

Cox proportional hazard model is developed by Cox (Cox, 1972) and has been used extensively in biomedical applications. Lane et al. (1986) first applied the proportional hazard model (PHM) to the prediction of bank failures. Such PHM has two advantages over the other classification techniques: first, it can be used to model the expected time to failure; secondly, it is a semi-parametric approach that does not impose any constraint on the distributional form of the hazard, thereby leaving the baseline hazard un-parameterized. The latter part is extremely useful when one is uncertain about the shape of the underlying hazard rate. Additionally, as suggested by Bartels (2003), the ability to justify a parametric model of social science theory can be a complicated task so that the Cox PHM can provide some flexibility over those parametric models.

The purpose of this paper is to present the application of an early warning model using the technique of Cox proportional hazard model to the bank failures of the late 2000s recession. The rest of the paper is organized as follows: the methodology section contains the general description of the Cox PHM; the data section describes the sample construction and data source; and the results, summary and conclusions are presented in the last few sections of this paper.

Methodology: The Cox Proportional Hazards Model

In this survival analysis, the central failure concept is the hazard rate, which defined as the probability that a bank will fail at time T given that it has survived through all of the previous time periods leading up to T. So T is the dependent variable in PHM that measures the time to failure for an individual bank. The survival function is defined as $S(\mathrm{t}) \equiv 1-\mathrm{F(t)}=\mathrm{P(T>t)}$, which represents the probability of surviving longer than t periods. The distribution function of time to failure is given by $F(\mathrm{t})=1-\mathrm{S(t)}$ with density function: $f(\mathrm{t})=-S'(\mathrm{t})$. The probability of leaving the initial state in the interval $[t, t+\Delta t)$ given survival up until time t is defined as $P(\mathrm{t} \leq \mathrm{T} < \mathrm{t}+\Delta\,\mathrm{t} \mid \mathrm{T} \geq \mathrm{t})$. The hazard function for T can then be expressed as:

$$h(\mathrm{t})=\lim_{\Delta t \to \infty} \frac{P(\mathrm{t} \leq \mathrm{T} < \mathrm{t}+\Delta\,\mathrm{t} \mid \mathrm{T} > \mathrm{t})}{\Delta t}=\frac{-S'(\mathrm{t})}{S(\mathrm{t})}$$

which specifies the instantaneous probability of failure under the assumption of survival up to time t (Wooldridge, 2002).

Unlike the parametric model that involves a certain distribution for hazard function (usually Log-logistic or Weibull distribution), the PHM proposed by Cox (1972) assumes the hazard function to follow a form given by:

$$h(\mathrm{t} \mid \mathrm{X}, \beta)=h_0(\mathrm{t})\exp\{\mathrm{X}'\beta\}$$

where X represents bank-specific variables that are assumed to affect the probability of failure, and β are the model coefficients to be estimated that can describe how each variable affects the likelihood of failure. The function $h_0(\mathrm{t})$ is a baseline hazard function, which is the hazard function of a bank with all X=0. Since all the explanatory variables are centralized, a bank with X=0 has values equal to the population mean. The function $h_0(\mathrm{t})$ has a nonparametric component that is assumed to be arbitrary and no distributional assumptions are imposed in the estimation of β or $h_0(\mathrm{t})$. β is a parametric vector that needs to be estimated. An exponential function is chosen for the parametric part since it can simplify the estimation of the regression coefficients (Cox and Oakes, 1984). Thus, Cox PHM is often considered as a semi-parametric function.

The survival function to be estimated here is:

$$S(\mathrm{t} \mid \mathrm{X}) = \mathrm{S}_0(\mathrm{t})^{\exp(\mathrm{X}'\beta)},$$

where $S_0(\mathrm{t}) = \exp[-\int_0^t h_0(\mathrm{u})\,\mathrm{du}]$.

Traditional proxies for the CAMELS[2] ratings have been validated as important determinants or predictors of bank failures in 2009 (Cole and White, 2012). Thus, for the parametric part of the function, we use covariates selected as proxies for the components of the CAMELS rating system. The estimating equation can thus be defined as follows:

$$X'\beta = \beta_0 AQCA_{it} + \beta_1 MR_{it} + \beta_2 PL_{it} + \beta_3 LPC_{it} + \beta_4 LPR_{it} + \beta_5 FA_{it} + \beta_6 Size_{it} + \beta_7 STECON_{it}$$

where: $AQCA_{it}$ are variables representing capital adequacy and asset quality; MR_{it} is a set of management risk variables; PL_{it} are variables that capture bank earnings (profitability) potential; LPC_{it} are variables that represent loan portfolio composition measures; LPR_{it} capture loan portfolio risk measures; FA_{it} are variables that represent funding arrangements; $Size_{it}$ is a structural factor variable, specifically representing bank size; $STECON_{it}$ are economic variables that capture macroeconomic conditions at the state level; t denotes the time when this duration analysis started.

[2] CAMELS rating system are used by regulators during on-site examinations to determine a bank's financial conditions, the letter stands for capital adequacy, asset quality, management quality, liquidity, and sensitivity to market risk as defined by FDIC.

Data Description

As in discriminate analysis, fitting the Cox PHM requires identifying the failed banks and a control group with a sample of survival banks. The data for both failed banks and surviving banks are collected from the Call Reports Database published on the website of Federal Reserve Board of Chicago (FRB). The banking data are available through the banks' quarterly financial statements made publicly available by the FRB. The 4th quarter Call Reports database is used to predict survival times during the period from first quarter of 2008 through the fourth quarter of 2012, since the late 2000s recession formally started in December 2007. The maximum survival time is censored at 21 quarters. The sample consists of all banks that failed between December 2007 and December 2012. Those banks that started their business operations after December 2007 were not included in the dataset to ensure the right censoring of data. Surviving or successful banks with missing values for any financial data being collected were discarded. Given these data restrictions, the resulting sample consists of 7,337 banks, of which contains 6,944 survival banks and 393 failed banks.

In addition to bank performance variables, this study also collected data from other sources that would reflect certain aspects of the local economic conditions during the recessionary period. These variables include state-level percentage change of monthly unemployment rate data that were obtained from the Bureau of Labor Statistics and were converted to quarterly data. State-level numbers of bankruptcy were collected from Bankruptcy filing statistics, published online by American Bankruptcy Institute (ABI). These bankruptcy figures were available for business, non-business and even sectoral (including agriculture-related filings under Chapter 12 bankruptcy) filings.

Estimation Results

The estimated results for Cox proportional hazard model are presented in table 2. Cox PHM estimates the coefficients associated with hazard rate, so a positive (negative) coefficient in the model indicates an increase (decrease) in the hazard rate as well as a decrease (increase) in the probability of survival.

Pursuant to the verified effectiveness of the loan portfolio diversification strategy, the loan portfolio composition variables identify the sectors that banks should consider in their loan servicing operations. The regression result indicates that banks may consider loan exposures to their consumer credit clientele (CONSTOTAL) before the onset of bank failures. This is consistent with the finding from Cole and Whitt (2012) who claimed that banks have comparative advantage in well-behaved consumer loans and thus consumer loan exposures should have negative impact on probability of failure. Similarly, agricultural (AGTOTAL) and industrial (INDUSTOTAL) loans are also negatively signed, which suggests that an increase in the portfolio of these loans will decrease the hazard rate, thus increase the probability of survival.

Loan portfolio risk variables (AGNR, AGR, CONSUM and INDUS) are calculated in ratio form as past due divided by nonaccrual loans. The most notable detection is the insignificance of both the non-real estate and real estate delinquency ratios for agricultural loans (AGNR and AGR) in the functional part that predicts the probability of survival. This suggests that agricultural loan delinquency ratios cannot be used as effective indicators for predicting bank survival. This finding confirms the statement from Ellinger and Sherrick (2011), as well as some recent literatures discussing U.S. agriculture during this 08-09 Recession (Li et al 2012, Sundell and Shane 2012). Agribusiness operations in general were doing well during the latest recession.

It has been stated that agricultural sector is in a strong position largely because of international trade. The large amount of export, which came from developing countries that facing an ever-growing demand for food, has provided enough cushions for agriculture to survive the recession. Thus, significant credit exposure to seemingly riskier agribusiness operations does not really pose as a risk or enhances a bank's tendency to fail.

On the contrary, the delinquency loan ratios for consumer loans (CONSUM) and commercial/industrial loans (INDUS) are significant positive regressors. These results may indicate that these two non-performing loan sectors may have significant adverse effects on the efficiency of banks and decrease the banks' survival probability.

The LOANHER is the Hirschman-Herfindahl Index (HHI) that is used to measure the loan diversification. The boundaries of the HHI are given by:

$$\frac{1}{n} \leq HHI \leq 1$$

where n stands for the loan segments. This index will approach to 1 if all loans are originated in one segment, so a high HHI is associated with higher probability of failure. Herfindahl index approach was included in Thomson's study and did not fare well in his regression models. In this model, however, this variable is positive as expected and significant under the 10 percent confidence level. The loan portfolio diversification is normally regarded as a risk-reducing strategy and, thus, the significant positive coefficient result in the model suggests that the "centralization" of loans indeed increases the probability of bank failure.

Variables that capture management risk and insider abuse are expected to be positively related to the hazard. However, the coefficient of insider loan (INSIDELN) is negative and not significant in contrast to the results obtained in previous studies. On the other hand, the overhead cost ratio (OVERHEAD) variable is negatively signed and significant. This contrasting result can be attributed to some plausible strategic moves of banks during the recessionary period. When faced with financial difficulty, especially illiquid conditions, banks may have the tendency to resolve the operating constraint by selling low-risk assets (like Treasury securities) that are relatively more easily marketable. As a result of such probable coping mechanism, the bank loses its asset base (denominator of the OVERHEAD ratio) while at the same time, overhead costs (the ratio's numerator) could possibly be rising as a result of higher degrees of operating inefficiency produced by less prudent operating decisions. Thus, the net effect of these two trends would be the positive relationship between increasing OVERHEAD ratios and the probability of bank failure.

The negative and significant coefficient on PROFIT indicates that the probability of failure is a negative function of earnings, which confirms the results of previous bank failure analysis.

Three variables are included to address interest rate risk, which is the sensitivity of all loans and deposits to relatively abrupt and unexpected shifts in interest rate. PURCHASEDTL, defined as the percentage of purchased liabilities among total liabilities, captures the rate-sensitive borrowing from such sources as federal funds. As described by Belongia and Gilbert (1990), the liabilities purchased from national market will have higher interest rates. Thus, according to theory, banks are more likely to fail when exposed to higher interest rate risk. However, this coefficient is negative and insignificant in this model. On the other hand, the coefficient for DEPLIAB is negative and significant. This result is consistent with the expectation that banks' tendency to thrive in their businesses is enhanced by their ability to maximize the generation of deposits to fund their business funding requirements. A third measure, duration GAP measurement, is also included in this analysis to further investigate interest rate risk issues. Duration GAP indicates the effect of interest rate changes on the net worth of the bank. The significant positive coefficient of GAP is consistent with logical expectations that higher GAP values are associated with higher interest rate risk. These results therefore imply that the probability of bank failure is positively related to the likelihood or incidence of higher interest rate risk or the banks' greater sensitivity to interest rate change.

The SIZE variable is significantly positively related to the probability of survival in the model. This result indicates that larger banks are more likely to fail during this recession and it is inconsistent with the finding discovered by Thomas (1991) on the "too big to fail doctrine", which mainly suggests that endangered or at-risk larger financial institutions will receive financial and other assistance from regulatory authorities because their failures are thought to impose severe damage to economy. However, this finding is not surprising since larger banks are heavily involved in the investment in RMBS (Cole and White, 2012), and evidence suggests that median assets and deposits of failed banks were considerably larger than non-failed banks in the latest recession (Aubuchon and Wheelock, 2010).

Two economic variables were included to reflect the local economic conditions. Percentage change of state-level unemployment rate (UNEMRATE) is expected to be positively related to the probability of bank failure for a deteriorating economic condition should have a negative impact on the banking industry. The coefficient in this model supports this macroeconomic concept with highly significant and positive sign. The state-level bankruptcy filing ratio (BF) variable's positive and significant coefficient result implies that a higher incidence of business or non-business failures or bankruptcies in each state would further depress the general economic conditions that would, in turn, influence the surge of bank failures.

Summary and Conclusion

This paper is designed to apply Cox proportional hazard model to identify the main bank-specific determinants of time to failure during the latest Great recession. The advantage of using Cox PHM is that its robust nature allows us to approximate the results for the correct parametric model when the underlying hazard function is unknown or in question. The other advantage Cox PHM over traditional classification techniques is that it can be used to study the relationship between survival time and bank-specific variables.

The covariates in the model include a set of variables that represents a bank's management decisions, operating strategies, financial conditions, and prevailing macroeconomic conditions. Non-performing loans, other than those associated with agricultural loans, have proven to have significant adverse effects on the banking system, and exposure to interest rate risk can also hurt the efficiency of bank performance.

All the covariates related to the banks' agricultural loan portfolios have insignificant impact on the survival probability. The healthy performance in agricultural sector is due to the benefits from strong balance sheets and sufficient exports to developing countries, so it is not surprising that even agricultural real and non-real estate loan delinquencies have not been established to significantly influence the likelihood of bank failure.

On the other hand, highly significant early warning signals are detected among consumer and commercial & industrial non-performing loans. As commercial/industrial loans are typically larger in magnitude, increases in delinquency in this loan category due to depressed economic demand and diminished economic activity will certainly help lead to bank failure.

Compared to larger banks, small banks are more fragile in the face of economic recession. However, the “too big to fail” doctrine has not been validated in this study as larger banks have been found to be more likely to fail in this recent recession, given their substantial exposure in the market for subprime mortgages.

Table 1. Definitions of Variables for the Duration Model

Variables	Descriptions
Dependent variable	
T	Length of time between t=1 and the subsequent failure date T
Explanatory variables	
AGNR	Aggregate past due/non-accrual agricultural non-real estate loans/total loans
AGR	Aggregate past due/non-accrual agricultural real estate loans/total loans
INDUS	Aggregate past due/non-accrual Commercial & Industrial loans /total loans
CONSUM	Aggregate past due/non-accrual Consumer loans /total loans
LOANHER	Loan portfolio Herfindahl index constructed from the following loan classifications: real estate loans, loans to depository institutions, loans to individuals, commercial & industrial loans, and agricultural loans.
AGTOTAL	Agricultural loans / total loans
CONSTOTAL	Consumer loans/total loans
INDUSTOTAL	Commercial & Industrial loans / total loans
OVERHEAD	Overhead costs/total assets
INSIDELN	Loans to insiders/total assets
PROFIT	Return on assets (Earnings)
SIZE	Natural logarithm of total assets
PURCHASEDTL	Purchased funds to total liabilities
DEPLIAB	Total deposits/ total liabilities
GAP	Duration GAP measure[3]
UNEMRATE	Percentage change of unemployment rate
BF	Business failure ratio

[3] GAP = Rate sensitive assets – Rate sensitive liabilities + Small longer-term deposits.

Table 2. Maximum Likelihood Result for Cox Proportional Hazard Model

Number of Observations: 7326

Log likelihood at convergence: -3038.7829

Variable	**Cox proportional hazards model**		
	Hazard coefficients	**Standard Error**	**t-statistics**
AGNR	-0.0168	0.0469	-0.36
AGR	-0.0012	0.0286	-0.04
INDUS	**0.0221**	**0.0055**	**4.06*** **
CONSUM	**0.0564**	**0.0194**	**2.91*** **
LOANHER	**1.2484**	**0.4626**	**2.70*** **
AGTOTAL	**-0.2110**	**0.1247**	**-1.69* **
CONSTOTAL	**-1.2533**	**0.2172**	**-5.77*** **
INDUSTOTAL	**-1.2969**	**0.5563**	**-2.33** **
OVERHEAD	**-0.8496**	**0.3732**	**-2.28** **
INSIDELN	-0.1358	0.2435	-0.56
PROFIT	**-0.8195**	**0.0963**	**-8.51*** **
SIZE	**0.4036**	**0.0499**	**8.08*** **
PURCHASEDTL	-0.0167	0.4112	-0.04
DEPLIAB	**-1.4732**	**0.4682**	**-3.15*** **
GAP	**0.3465**	**0.0214**	**16.18*** **
BF	**0.4973**	**0.2218**	**2.24** **
UNEMRATE	**0.8279**	**0.1541**	**5.37*** **

*** Denotes statistical significance at 1% level.

** Denotes statistical significance at 5% level.

* Denotes statistical significance at 10% level.

Reference

Aubuchon, Craig P. and David C. Wheelock, 2010, The geographic distribution and characteristics of US bank failures, 2007-2010: Do bank failures still reflect local economic conditions?, *Federal Reserve Bank of St. Louis Review* 92(05), 395-415.

Bartels, B.2003. "Advances in Duration Modeling: The Split Population Duration Model". Lab Notes,3,4-7. http://polisci.osu.edu/sites/polisci.osu.edu/files/AdvancesinDuration.pdf (Accessed December 2013)

Cox, D.R. 1972. "Regression models and life-tables." *Journal of Royal Statistical Society. Series B(Methodological)* 34(2): 187-220.

Cox, D.R. and D. Oakes, 1985. Analysis of survival data (Chapman and Hall, London).

Cole, R.A., and J.W. Gunther. 1995. "Separating the likelihood and timing of bank failure." *Journal of Banking & Finance* 19(6):1073-1089.

Cole, R.A., White, L.J., 2012. "De´ja vu all over again: the causes of US commercial bank failures this time around." *Journal of Financial Services Research* 42, 5–29.

Ellinger, P. and B. Sherrick, *Financial markets in agriculture*. Illinois: Illinois Farm Economics Update, October 15,2010.

http://www.farmdoc.illinois.edu/ifeu/IFEU_08_02/IFEU_08_02.html

(Accessed December 2013)

Hanweck, G.A. 1977. "Predicting bank failure." *Research Papers in Banking and Financial Economics*.

Hoggarth, G., Reis, R., and Saporta, V., 2002. “Costs of banking system instability: some empirical evidence.” *Journal of Banking and Finance* 26, 825-855

Lane, W.R., S.W. Looney and J.W. Wansley, 1986. “An application of the Cox proportional hazards model to bank failure”, *Journal of Banking and Finance* 10, 511-531

Li, X., C. Escalante, J.E. Epperson, and L.F. Gunter. 2013. "Agricultural Banking and Early Warning Models of Bank Failures for the Late 2000s Great Recession." *Agricultural Finance Review* 73(1):8-8.

Martin, D. 1977. "Early warning of bank failure: A logit regression approach." *Journal of Banking & Finance* 1(3):249-276.

National Bureau of Economic Research (NBER), *Announcement made by Business Cycle Dating Committee*, Cambridge,MA. Internet site:

http://www.nber.org/cycles/sept2010.html (Accessed December 2013)

Schmidt, P., and A.D. Whitte, 1989. “Predicting criminal recidvism using ‘split population’ survival time models.” *Journal of Econometrics* 40, 141-159.

Sundell, P., and M. Shane, 2012. “The 2008-09 Recession and Recovery: Implications for the Growth and Financial Health of U.S. Agriculture.” United States Department of Agriculture, Economic Research Service (USDA, ERS) Online publication.

http://www.ers.usda.gov/publications/wrs-international-agriculture-and-trade-outlook/wrs1201.aspx#.UqZ0-tJDu7w (Accessed December 2013)

Thomson, J.B. 1991. "Predicting bank failures in the 1980s." *Economic Review* 27:9-20.

Wheelock, D.C. and Wilson, P.W. 1995. "Explaining bank failures: deposit insurance, regulation, and efficiency." *Review of Economics and Statistics*, 77, 689-700.

Wooldridge, J.M. 2002. *Econometric analysis of cross section and panel data*: The MIT press.