Article:

Alzugaiby, Basim, Gupta, Jairaj, Mullineux, Andrew et al. (2021) Relevance of size in predicting bank failures. International Journal of Finance and Economics. pp. 3504-3543. ISSN: 1099-1158

https://doi.org/10.1002/ijfe.1973

Received: 7 December 2018 | Revised: 15 January 2020 | Accepted: 18 June 2020
DOI: 10.1002/ijfe.1973

RESEARCH ARTICLE

WILEY

Relevance of size in predicting bank failures

Basim Alzugaiby[1] | **Jairaj Gupta**[2] | **Andrew Mullineux**[2] | **Rizwan Ahmed**[2]

[1]College of Economics and Administrative Sciences, Al Imam Mohammad Ibn Saud Islamic University, Riyadh, PO Box 5701, Saudi Arabia

[2]Department of Finance, University of Birmingham, Birmingham, UK

Correspondence
Jairaj Gupta, Department of Finance, University of Birmingham, Birmingham B15 2TY, UK.
Email: J.Gupta@outlook.com

Abstract

Employing a statistical model-building strategy, this study aims to analyse the United States' bank failures across different size categories (small, medium, and large). Our results suggest that factors associated with bank failures vary across respective size categories, and the average marginal effects (AMEs) of mutually significant covariates also exhibit significant variability across different size classes of banks. The results are robust to up-to 3 years of lagged regression estimates, various control variables, interaction between bank size and bank charter, alternative bank size classifications, and macroeconomic crisis periods.

KEYWORDS

bank failures, bank size, banking, default risk, systemic risk

JEL CLASSIFICATION

G01; G21; G28

1 | INTRODUCTION

Does size matter in predicting bank failures? The answer to this question would be helpful to policymakers and bank regulators seeking to improve their understanding of bank failures across different size categories, and thereby promoting enhanced stability of the financial system. This issue was seriously exaggerated following the failure of large and complex banks in 2008 (the recent financial crisis) which resulted in extremely high costs to national economies, as they were forced to bail them out in order to restore confidence in the financial markets (Pais & Stork, 2013). Over the last three decades, several banks have been criticized for becoming oversized and thereby carrying the associated higher systemic risk. In response, several restrictions have been enacted by federal governments to downsize or split up these banks to reduce the public finance risk. For instance, the Dodd-Frank Act of 2010 is a United States (US) federal law intended to limit banks' involvement in some risky activities and to ban mergers that result in a financial institution with total liabilities surpassing 10% of the aggregate consolidated liabilities of all financial firms (to prevent the emergence of "too big to fail" banks [Bertay, Demirgüç-Kunt, & Huizinga, 2013]). Proponents of the act also argue that the constraints, particularly size limitation, shall prevent future crises and protect consumers from abusive financial services practices. However, many argue that these actions would impair the efficiency of capital allocation for some banks and add costs to the economy (Aiyar, Calomiris, & Wieladek, 2014). Others also argue that such restrictive regulations may lead to the failure of many small banks deemed "too important to fail", which may cause the recurrence of financial crises (De Haan & Poghosyan, 2012). This debate reveals the need for further investigation into the heterogeneity of bank failures across different size classes, to recognize the similarities and differences before taking appropriate measures.

The literature on individual bank failures is extensive and offers a rich assessment of several aspects of bank failures (e.g., Berger, Imbierowicz, & Rauch, 2016; Cole & White, 2012; DeYoung & Torna, 2013; Kolari, Glennon, Shin, & Caputo, 2002; Lane, Looney, & Wansley, 1986; Meyer & Pifer, 1970; Schaeck, 2008; Thomson, 1992; Wheelock & Wilson, 2000). However, the factors and the extent to which they related to the probability of bank failures across size classes remain mostly overlooked.[1] This is perhaps surprising because the literature shows that bank size is as an essential economical foundation as the capital (Berger & Bouwman, 2013), and plays a crucial role in many dimensions such as performance (e.g., Bertay et al., 2013), financial stability (e.g., Bhagat, Bolton, & Lu, 2015), scope (e.g., De Jonghe, Diepstraten, & Schepens, 2015), lending (e.g., De Haas, Ferreira, & Taci, 2010), funding strategies (e.g., Loutskina, 2011), and systemic risk (e.g., Laeven, Ratnovski, & Tong, 2016). Notwithstanding the evidence of bank size heterogeneity effects on various aspects, particularly financial stability, the literature lacks a thorough analysis of determinants and predictability of bank failures across bank size categories.

Considering the discussion above, the aim of this study is to empirically analyse whether and how banks' failure predictors vary across different size categories (using banks' total assets in a given year t to classify banks into small, medium, and large banks). Using panel logistic regression technique, we develop separate failure prediction models for small, medium, and large banks in the US (from 1985 until 2016) and report any differences in comparison to an all-size inclusive failure prediction model. Another contribution of this paper is that, unlike previous studies that draws heavily on accounting-based predictors such as capital, earnings, and liquidity ratios (e.g., Berger et al., 2016; Cole & Gunther, 1995; Cole & White, 2012; Kolari et al., 2002); our study follows the suggestions of Gupta, Gregoriou, and Ebrahimi (2018) and statistically analyses the relative importance of a comprehensive set of predictors (found significant in prior bank failures literature) to develop parsimonious multivariate failure prediction models.

Thus, in the first step we employ univariate regression analysis as a variable selection technique to investigate the relative importance of numerous accounting-based variables used in previous bank failure literature. Specifically, the broad categories of CAMELS, where the letters refer to Capital adequacy (e.g., total equity to total assets ratio), Asset quality (e.g., non-performing loans to total assets ratio), Management quality (e.g., cost to income ratio), Earnings (e.g., net interest margin), Liquidity (e.g., cash and due to total assets ratio), and Sensitivity to market risk (e.g., trading income to total operating income ratio); and other categories such as funding, business model, and growth, are analyzed. We investigate a total of 61 accounting variables. Univariate regression analysis shows that average marginal effects (AMEs) of most accounting-based predictors used in the literature vary across size categories and across three respective lagged periods. Generally, the AMEs of respective covariates (1-year lag) for small banks are higher compared to estimates obtained for medium and large banks. However, for 2- and 3-years lagged estimates, AMEs of large banks are mostly the highest. This suggests that the prediction power of variables for small banks are stronger on short term, while for large banks it's stronger for longer horizon forecast. To narrow down the list of covariates for further multivariate analysis, we select only those variables that are significant in all three lagged univariate regression estimates. We repeat this process for small, medium, and large banks respectively. The final lists of variables for all, small, medium, and large banks contain 19, 19, 20, and 21 variables respectively.

Subsequently, following the multivariate model building strategy suggested by Gupta et al. (2018), we rank competing variables based on the magnitude of their AMEs, and then introduce each variable at a time in descending order of magnitude. We perform this to develop multivariate models for all, small, medium, and large-sized banks respectively. The rationale for this approach is that a variable with a higher value of AME induces higher change in the failure probability, and thus should be given priority in the variable selection process (Gupta et al., 2018). We exclude a variable from the multivariate models if, when added: (a) it changes the sign of any previously added variable; (b) it shows the opposite sign to that generated in univariate regression; (c) it holds the identical sign to univariate analysis, but is insignificant with a p-value greater than 0.10; or (d) it makes a previously introduced variable insignificant with a p-value greater than 0.10. We end up with varying sets of covariates with six, seven, six, and five variables (main variables) for multivariate models for all, small, medium, and large banks, respectively. Multivariate empirical results show that factors associated with bank failures and the magnitudes of mutually significant factors (Average Marginal Effects) vary across small, medium, and large size categories.

These results are robust to the presence of control variables including house price inflation, foreign ownership, and dummies for banking crises and regulators. We perform an additional robustness test by disaggregating our sample by banking crises, market crises, and normal times, and treating them as separate groups. We also use an alternative size classification. We rerun all multivariate regressions separately for all, small, medium, and large banks, and qualitatively similar results are obtained.

Our findings emphasize the importance of considering bank size when designing appropriate policies and regulations targeted toward enhancing financial stability and resilience. Future studies should, whenever possible, separate banks by size category to clearly understand the heterogeneity in bank failures.

The remainder of this paper is structured as follows. In Section 2, we provide a review of literature on of bank failures and research objective. Section 3 presents discussion on the dataset, sample, and covariates. In Section 4, we outline empirical methods and discuss our results. Sections 5 presents additional analysis and robustness test. Section 6 concludes this study.

2 | LITERATURE REVIEW AND RESEARCH OBJECTIVE

This paper is guided by theoretical models and empirical literature related to determinants and prediction of bank failures. In theory, two views explain the sources of bank defaults. First is the panic-based view introduced by Bryant (1980), which posits that banks are inherently vulnerable and are subjected to contagion (Calomiris, 2007). According to this view, bank runs can be attributed to the strong likelihood of depositors withdrawing their funds because others will run, or due to ambiguous or inaccurate information about the institution's health (Diamond & Dybvig, 1983). In such circumstances, many banks fail due to high withdrawal pressure and risk spreading the adverse effects within the banking system, including solvent banks. Second is the fundamental-based view which considers banks to be inherently stable and not vulnerable to panic. According to this view, depositors withdraw their funds due to adverse fundamental changes in the economic conditions of banks (e.g., large losses), leading to the failure of only weak and fragile banks (Calomiris, 2007). The latter view supports our paper, which aims to investigate bank default predictors. We believe that the financial status of a bank generally governs current depositors' withdrawal decisions, investors, and expected depositors. Thus, it is essential to focus on the factors that determine the financial condition of banks, in order to assist interested parties in making informed decisions.

The empirical literature on the determinants of bank failures typically concentrates on the United States (US) banks and thrifts (e.g., Berger et al., 2016; Cole & White, 2012; Lane et al., 1986; Meyer & Pifer, 1970; Schaeck, 2008; Thomson, 1992; Wheelock & Wilson, 2000). Furthermore, the literature draws heavily on accounting-based indicators and aims to construct early warning models generally based on the Uniform Financial Rating System, informally known as the CAMELS ratings system, to identify distress institutions prior to their failures (e.g., Cole & Gunther, 1995; Cole & White, 2012; Kolari et al., 2002). Several studies supplement the CAMELS proxies with some information about audit quality (Jin, Kanagaretnam, & Lobo, 2011), or corporate governance (ownership, management, and compensation) (Berger et al., 2016). All of these studies show that their models are significant and effective in predicting bank failures. Also, several statistical (e.g., Discriminant analysis, DA, Logit/Probit regression models) and intelligence (e.g., Support Vector Machines, SVM, Neural Networks) techniques have been used to analyse and predict bank failures. Demyanyk and Hasan (2010) provide a thorough review of these techniques and related studies; we refer interested readers to this study for more details.

The vast body of research focuses on bank failures that occurred during either the saving and loan crisis period of 1987–1992, or the 2008–2010 subprime lending crisis period. Papers studying the failed banks during the saving and loan crisis (e.g., Cole & Gunther, 1995; DeYoung, 2003; Wheelock & Wilson, 2000) show that banks with poor capitalization, extreme non-performing loans, low earnings, and less liquidity were associated with a higher probability of failure. Recently, several studies analysed the determinants of bank failures in the United States during the recent subprime lending crisis (Berger et al., 2016; Cole & White, 2012; DeYoung & Torna, 2013; Hong, Huang, & Wu, 2014; Imbierowicz & Rauch, 2014; Ng & Roychowdhury, 2014). Cole and White (2012) use the CAMELS indicators together with measures of "traditional" banking activities, such as commercial and residential loans, to explain the drivers of US commercial bank failures that occurred between 2004 and 2008, and to predict 2009 failures. They find that banks with less capital, bad asset quality, lower earnings, less liquidity, and with higher loan allocations to construction-and-development loans, commercial mortgages, and multi-family mortgages, are more likely to fail. DeYoung and Torna (2013) focus on "non-traditional" banking activities with mainly noninterest income such as stakeholder activities and Fee-for-Service income to analyse the US bank failures from 2007 to 2009. They find that stakeholder activities (e.g., investment banking, insurance underwriting, proprietary trading, and venture capital) increase the probability of bank failures only if the bank was already suffering from financial distress, whereas Fee-for-Service income (e.g., insurance sales, loan servicing and securities brokerage) reduce the probability of bank failures during the crisis. Hong et al. (2014) examine the links between US commercial bank failures and Basel III liquidity risk measures, liquidity coverage ratio (LCR), and net stable funding ratio (NSFR). They report that both LCR and NSFR have limited effects on explaining bank failures. Testing the impact of loan loss reserves on US bank failures, Ng and Roychowdhury (2014) employ a Cox proportional-hazard model and report that "add-backs" of loan loss reserves is positively related to bank failures. Additionally, Imbierowicz and Rauch (2014) investigate the impact of liquidity risk and credit risk on probabilities of default in US commercial banks. They document that these two risk sources separately increase the likelihood of default, but their joint effect can either aggravate or mitigate default risk. More recently, Berger et al. (2016) analyse the roles of corporate governance (ownership, management, and compensation structures) in US commercial bank failures. They find that banks with more shareholdings of lower-level managers and non-CEO higher-level managers are more likely to fail. However, the shareholdings of CEOs do not increase the risk of failure.

According to Berger and Bouwman (2013), the existing literature generally suffers from two respective limitations. First, most studies cover a short period of time (the span of one banking crisis) and do not pay attention to the periods prior to and following the crisis (normal times), or other banking crises. Second, a thorough analysis of bank failures across size classes is largely ignored. This suggests that the findings of studies reviewing the saving and loans crisis consider small banks results, given the domination of failures among small banks (Berger & Bouwman, 2013). Thus, the primary objective of our study is to empirically examine whether and how the failure predictors show a discrepancy among virtually all US commercial bank failures in different size classes.

The existing literature indicates that bank size plays a pivotal role in maintaining financial stability (e.g., Bhagat et al., 2015; De Haan & Poghosyan, 2012; De Nicolo, 2000; Demsetz & Strahan, 1997). Demsetz and Strahan (1997) focus on US bank holding companies (BHCs) to analyse the relationship between bank size and volatility in stock prices as a measure of risk. They conclude that large BHCs are better diversified, but they are not less risky than small BHCs. Analyzing an international sample of banks, including 419 BHCs in the US, De Nicolo (2000) finds a positive relationship between bank size and volatility in small to medium-sized BHCs and a negative relationship in large ones. Hakenes and Schnabel (2011) analyse the relationship between bank size and risk-taking under the Basel II Capital Accord. They conclude that large banks have an advantage over small banks to choose between the Standardized and Internal Ratings Based Approach which pushes small banks to take more risk. Moreover, De Haan and Poghosyan (2012) report a non-linear relationship between size and earnings volatility. They find that bank size is negatively related to earnings volatility, but the relationship becomes positive when a bank's total assets exceed $5 billion. Recently, Bhagat et al. (2015) studied the size effect on the risk-taking of US based financial institutions, including commercial banks, investment banks and life insurance companies. They document a positive relationship between bank size and risk in the pre-crisis period (2002–2006) and the crisis period (2007–2009), but not in the post-crisis period (2010–2012). Overall, these analyses provide useful insights that contribute to the main objective we formulate in this paper.

To the best of our knowledge, only two papers can be considered as closely related studies to our paper (Berger & Bouwman, 2013 and Imbierowicz & Rauch, 2014). Berger and Bouwman (2013) examine the impact of capital on bank performance (survival and market share) across bank size classes (small, medium, and large), and how this effect differs across banking crises, market crises, and normal times between 1984 and 2010, in the United States. They find that capital improves the performance of medium and large banks only during banking crises and helps to improve the performance of small banks during banking crises, market crises, and normal times. However, Berger and Bouwman's paper differs from ours in many respects. First, their study is based on only one of the six CAMELS components (capital), and ignores the others, that may misclassify distressed banks (Cole & White, 2012). Second, they use a development sample up to 2010, while we extend our sample to cover the most recent observation (i.e., up to 2016). Third, they split the bank size classes into small banks (gross total assets, or GTA, up to $1 billion), medium banks (GTA exceeding $1 billion and up to $3 billion), and large banks (GTA exceeding $3 billion), while we use a different and arguably more accurate criteria to determine bank size.[2] Specifically, in any given year *t*, banks corresponding to the bottom 25 percentile of total assets are considered small banks, the top 25 percentile are considered large banks, and the rest are medium banks. Fourth, they exclude banks that are below $25 million of total assets instead of including all banks, as we do.

Finally, Imbierowicz and Rauch (2014) investigate the relationship between liquidity risk and credit risk in different bank size categories. They show that liquidity risk is slightly larger for small and medium sized banks. However, they have not focused on the discrepancies in the determinants of bank failures across size categories.

3 | DATASET, SAMPLE AND COVARIATES

The data used in our empirical analysis come from the Federal Deposit Insurance Corporation (FDIC) database. The FDIC collects financial information such as balance sheets and income statements from the Consolidated Reports of Condition and Income (Call Reports) submitted by US financial institutions on a quarterly basis. In line with several existing studies we focus only on commercial banks to obtain a homogenous sample. We exclude savings banks due to the discrepancy in directions between these banks and the commercial banks (Cole & White, 2012). To construct financial variables, we use the year end (fourth quarter) data from 1985 to 2016 for each bank in our sample.

3.1 | Defining bank failures

To identify commercial bank failures, we use the Failed Bank list reported by the FDIC, which is widely used in the existing literature (e.g., Berger et al., 2016; Liu & Ngo, 2014). The list contains characteristics of failed banks, including bank names, locations, acquiring institutions, and closing dates. The FDIC generally records a bank as failed if it enters either "assistance transactions," which require restructuring and the charter survives, or "outright failure," in which a bank closes its operations and the charter is terminated. The failure list in our sample contains 1,871 banks with 1,694 outright failures and 123 assistance transactions.

3.2 | Defining small, medium, and large banks

The literature documents the importance of bank size and the advantages generated by size heterogeneity (e.g., Berger & Bouwman, 2009, 2013). However, there is no formal definition that identifies bank size classes. Thus, we use similar criteria that have applied by Imbierowicz and Rauch (2014), which is based on the percentile of bank's total assets in a given year to classify it as small, medium, or large. Specifically, we consider banks corresponding to the bottom 25 percentile of total assets as small banks, the top 25 percentile as large banks, and the rest as medium banks. We perform this exercise on a yearly basis, as our size classification is based on the relative assets size of respective banks, which changes from 1 year to another due to various reasons. This gives us a sample of 74,533 bank-year observations for small banks, 149,072 bank-year observations for medium banks, and 74,520 bank-year observations for large banks. This subsequently leads to 8,260 small banks, 12,977 medium banks, and 7,210 large banks in our sample.[3]

3.3 | Sample description

Table 1 presents the annual failure rates of banks from 1985 to 2016. To observe any differences between size categories, we also report the annual failure rates of small, medium, and large banks. The average failure rate of our entire sample is around 0.54%. The average failure rate is highest for small banks (0.67%), followed by large banks (0.53%), and lowest for medium banks (0.47%). Further, we see in Table 1 that the relationship between failure rate and bank size is most likely negative, up until the onset of the subprime lending crisis in 2008. This relationship turns out to be positive, specifically between 2008 and 2012. However, after the crisis period, it becomes negative.

The failure rates of small banks experienced a significant rise around the savings and loan crisis of the 1980s and 1990s, followed by large banks, and were lowest for medium banks. Yet the failure rates of large banks escalated dramatically during the subprime lending crisis, followed by medium banks, and were lowest for small banks (see Figure 1). This transformation in the failure rates may be attributed to the augmentation in bank size associated with high risk-taking by these banks (due to the moral hazard that the government will bail them out in troubled times to stabilize the financial system and avoid unfortunate consequences to the economy [Pais & Stork, 2013]).

3.4 | Covariates

In this section, we discuss the rationale behind our choice of dependent variable, followed by relevant discussion on the explanatory and control variables employed in this study (see Table A1 in the appendix).

3.4.1 | Dependent variable

One important focus of this study is the determination of factors that associated to bank failures across different size classes. Therefore, the dependent variable is binary (fail/non-fail). As discussed in Section 3.1 and following Liu and Ngo (2014), we consider all banks in the FDIC failed list as failed banks if presented as either "assistance transactions" or "outright failures".

TABLE 1 Failures rate of US banks

Year (1)	All banks Failures (2)	Total (3)	% failures (4)	Small banks Failures (5)	Total (6)	% failures (7)	Medium banks Failures (8)	Total (9)	% failures (10)	Large banks Failures (11)	Total (12)	% failures (13)
1985	118	14,656	0.8051	54	3,664	1.4738	56	7,328	0.7642	8	3,664	0.2183
1986	142	14,468	0.9815	61	3,616	1.6869	63	7,235	0.8708	18	3,617	0.4977
1987	199	14,171	1.4043	107	3,542	3.0209	73	7,086	1.0302	19	3,543	0.5363
1988	276	13,626	2.0255	86	3,406	2.5250	113	6,813	1.6586	77	3,407	2.2601
1989	204	13,074	1.5603	80	3,268	2.4480	82	6,538	1.2542	42	3,268	1.2852
1990	158	12,643	1.2497	70	3,161	2.2145	67	6,321	1.0600	21	3,161	0.6643
1991	103	12,258	0.8403	34	3,063	1.1100	46	6,132	0.7502	23	3,063	0.7509
1992	75	11,796	0.6358	26	2,948	0.8820	35	5,898	0.5934	14	2,950	0.4746
1993	38	11,303	0.3362	13	2,826	0.4600	17	5,651	0.3008	8	2,826	0.2831
1994	11	10,820	0.1017	3	2,705	0.1109	3	5,410	0.0555	5	2,705	0.1848
1995	5	10,271	0.0487	1	2,567	0.0390	1	5,137	0.0195	3	2,567	0.1169
1996	4	9,897	0.0404	1	2,474	0.0404	3	4,949	0.0606	0	2,474	0.0000
1997	1	9,562	0.0105	1	2,391	0.0418	0	4,781	0.0000	0	2,390	0.0000
1998	3	9,131	0.0329	1	2,283	0.0438	1	4,566	0.0219	1	2,282	0.0438
1999	6	8,838	0.0679	3	2,210	0.1357	2	4,419	0.0453	1	2,209	0.0453
2000	6	8,597	0.0698	2	2,150	0.0930	4	4,298	0.0931	0	2,149	0.0000
2001	3	8,284	0.0362	3	2,071	0.1449	0	4,142	0.0000	0	2,071	0.0000
2002	10	8,035	0.1245	4	2,009	0.1991	3	4,018	0.0747	3	2,008	0.1494
2003	1	7,896	0.0127	1	1,975	0.0506	0	3,948	0.0000	0	1,973	0.0000
2004	3	7,760	0.0387	1	1,941	0.0515	2	3,879	0.0516	0	1,940	0.0000
2005	0	7,671	0.0000	0	1,918	0.0000	0	3,836	0.0000	0	1,917	0.0000
2006	0	7,568	0.0000	0	1,892	0.0000	0	3,784	0.0000	0	1,892	0.0000
2007	1	7,444	0.0134	0	1,861	0.0000	1	3,722	0.0269	0	1,861	0.0000
2008	22	7,238	0.3040	4	1,810	0.2210	5	3,619	0.1382	13	1,809	0.7186
2009	124	7,018	1.7669	11	1,755	0.6268	57	3,509	1.6244	56	1,754	3.1927
2010	130	6,765	1.9217	19	1,692	1.1229	55	3,383	1.6258	56	1,690	3.3136
2011	84	6,443	1.3037	9	1,611	0.5587	51	3,222	1.5829	24	1,610	1.4907
2012	40	6,235	0.6415	8	1,559	0.5131	26	3,118	0.8339	6	1,558	0.3851
2013	23	5,999	0.3834	11	1,500	0.7333	9	3,000	0.3000	3	1,499	0.2001

(Continues)

TABLE 1 (Continued)

Year (1)	All banks Failures (2)	Total (3)	% failures (4)	Small banks Failures (5)	Total (6)	% failures (7)	Medium banks Failures (8)	Total (9)	% failures (10)	Large banks Failures (11)	Total (12)	% failures (13)
2014	14	6,532	0.2143	7	1,633	0.4287	5	3,266	0.1531	2	1,633	0.1225
2015	8	6,199	0.1291	5	1,550	0.3226	2	3,100	0.0645	1	1,549	0.0646
2016	5	5,927	0.0844	4	1,482	0.2699	1	2,964	0.0337	0	1,481	0.0000
Average			0.5370			0.6740			0.4715			0.5312

Notes: The table reports annual details of failed and non-failed US commercial banks. Column 1 lists years followed by the number of failed banks in that year (column 2), total number of banks in the database in that year (column 3), and percentage of failed banks (failed/Total banks × 100) in that year (column 4) for our entire sample of banks. The following columns show identical information for small, medium, and large sized banks. In the last row, "Average" is the mean of annual failure rates reported in columns 4, 7, 10 and 13 respectively.

3.4.2 | Explanatory variables

To develop our multivariate regression models, we consider a broad list of 61 financial (accounting-based) variables as candidate failure predictors, and briefly explain them in the Table A1. These predictive variables are drawn from popular studies on bank failures, including Wheelock and Wilson (2000), Kolari et al. (2002), Arena (2008), Cole and White (2012), DeYoung and Torna (2013), Betz, Opriсă, Peltonen, and Sarlin (2014), and many others.[4] We do not consider market-based covariates for two reasons. First, the vast majority of our sample comprises unlisted banks. Second, our prediction horizon is 1 to 3 years prior to failure, while the signals of these variables tend to have a shorter-run time horizon (Betz et al., 2014). Moreover, Cole and Wu (2009) suggest that bank-specific variables are more essential than market and macroeconomic variables when predicting bank failures. Our choice of variables reflects all dimensions in the CAMELS categories, as well as funding, business model, leverage, off-balance sheet, growth, non-traditional activities, and others.[5]

Capital adequacy

Capital is the most important indicator that is considered in all regulator and supervisor frameworks (e.g., Basel) to ensure the safety and soundness of banks and financial systems. It is also included as a key variable in virtually all previous studies. The level of capital reflects the capacity of banks to meet their financial obligations. Hence, a decline in capital is a clear sign of potential financial troubles. To measure the capital adequacy we use the total equity to total assets (TETA) ratio, which is largely used in the literature and a highly valuable proxy of capital, and the nonperforming assets coverage ratio (NPACR), which is shown by Chernykh and Cole (2015) to outperform regulatory capital ratios in predicting US bank failures. Higher values of these indicators are expected to reduce the probability of bank failures. Following Poghosyan and Čihak (2011), we do not incorporate the ratio of regulatory capital to total risk-weighted assets to avoid any risk assessment, and because the calculation of these ratios is based on relatively arbitrary weights.

Asset quality

poor quality of assets generally increases the probability of bank failures. The most preponderant and risky assets of commercial banks are loans. Thus, we focus heavily on this asset group and employ a wide variety of potential indicators, specifically loan loss reserves, loan loss provisions, net charge off, and all types of non-performing

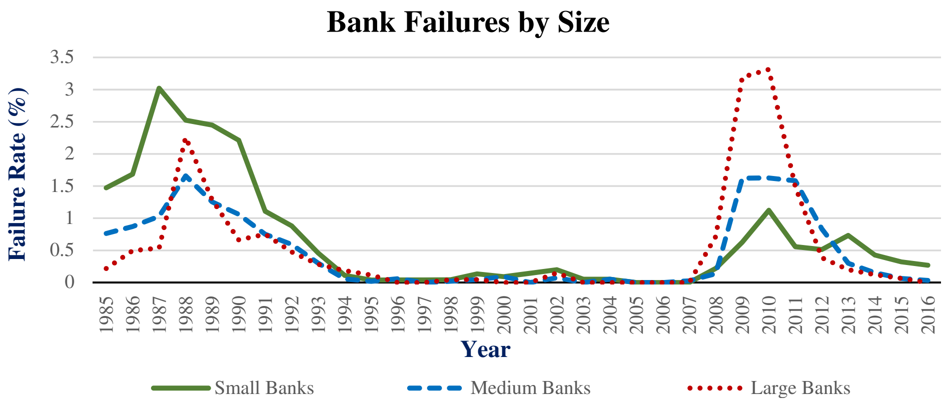


FIGURE 1 This figure shows the failure rate (in %) across bank size categories that occurred during our sample period from 1985 until 2016 [Colour figure can be viewed at wileyonlinelibrary.com]

loans. In general, these variables are expected to have a positive relationship with bank failures probability.

Management

Management competence plays a central role in the performance and success of a bank. Although the management quality is difficult to measure with financial data, DeYoung (1998) documents that cost efficiency reflects management quality. He concludes that higher management quality leads to higher efficiency of resource uses, thus we use the cost efficiency represented by cost-to-income ratio to gauge the quality of management. Following DeYoung and Torna (2013) we also use cost inefficiency, measured by total noninterest expenses to total assets. These indicators are expected to be positively associated with bank failures.

Earnings

This category reflects the profitability and performance of banks. The most frequently applied measures are return on assets (ROA), return on equity (ROE), and net interest margin (NIM). Higher earnings enhance the profitability (ROA, ROE, NIM) and capital level (equity/assets) that lead to improved bank performance. Hence, the relationship between profitability and the probability of bank failures is expected to be negative.

Liquidity

An adequate liquidity is essential for banks to meet their current obligations and to cope with unexpected withdrawals of depositors without liquidating assets. To gauge this category, we employ most of the variables that have been used in the literature, including federal funds to total assets, securities to total assets, total loans to total deposits, and others (see Table A1). In general, we expect a higher value of these ratios to have a negative relationship with bank failures probability.

Sensitivity to market risk

This category is represented by the share of trading income (TIOI). Higher trading income could be associated with a riskier business model and higher probability of failing. Liquid, however, rather than loans, is more likely to decrease fire sale losses. Thus, it is difficult to predict the direction of the influence in advance.

In addition to the CAMELS covariates, we also include many other potential explanatory variables, specifically to measure funding, business model, leverage, off balance sheet, growth, non-traditional activities, and others (see Table A1).

3.4.3 | Control variables

To establish the robustness of our explanatory variables, we also report our multivariate results, supplementing the following control variables:

Primary regulator

US commercial banks are regulated by one of three federal regulators. National banks are regulated by the Comptroller of the Currency (OCC), state-chartered banks that are members of the Federal Reserve System (FRS) are regulated by the Federal Reserve, and state-chartered banks that are not members of the FRS are regulated by the Federal Deposit Insurance Corporation (FDIC). To investigate the influence of the regulator on bank failures, we include three dummies: OCC, FED, and FDIC. Due to collinearity, we use only two of them (FED and FDIC), and treat OCC as the reference category.

Foreign ownership

Foreign ownership is captured by a dummy variable that takes the value of 1 if 25% or more of a bank is foreign-owned, and 0 otherwise. Arena (2008) concludes that foreign banks in emerging countries can mitigate their probability of failure due to better risk-based management practices, capitalization, and access to parent funding, however in the United States, Berger, DeYoung, Genay, and Udell (2000) find that domestic banks are generally more efficient than foreign banks. We therefore expect a positive relationship between foreign ownership and the probability of failure.

Growth of house prices index

This economic variable is a broad measure to capture real estate prices at state-level. The movements of the real estate prices can impair the stability of banks because defaulted mortgage loans are generally covered by real estate as collaterals, and banks will not be able to recover all of the value of collaterals in a situation of deteriorating real estate prices. To capture the effect of this variable, we obtain the seasonally adjusted house price indices (HPIs) from the Federal Housing Finance Agency. Following Berger and Bouwman (2013) we use all transactions index (based on purchases and appraisals) data until 1990 and purchase only index (based on purchases) data from 1991 onward.

Banking crises

To measure the effects of previous banking crises, we create two dummy variables. First, the saving and loans crisis that takes the value of 1 for the years from 1987 to 1990, and 0 otherwise. Second, the subprime lending crisis takes the value of 1 for the years from 2008 to 2010, and 0 otherwise.

4 | ECONOMETRIC METHOD AND ANALYSIS

This section discusses the statistical technique employed in this study followed by multivariate model building strategy.[6] Subsequently it presents our empirical results and analysis.

4.1 | Panel logistic regression

Numerous statistical methodologies have been used to analyse and predict bank failures. These methods range from simple Discriminant Analysis (e.g., Haslem, Scheraga, & Bedingfield, 1992) and Logit/Probit regressions (e.g., Berger et al., 2016) to advanced machine learning techniques, such as Extreme Gradient Boosting (e.g., Climent, Momparler, & Carmona, 2019). To investigate the factors that associated with bank failures and establish our empirical validation, we use panel logistic regression with random effects. Although hazard models are emerging as a popular choice (e.g., Cole & Wu, 2009; Ng & Roychowdhury, 2014), Gupta et al. (2018) argue that the discrete-time hazard model with logit link is essentially a panel logistic model that controls for firms' age. Accordingly, we assume that the marginal probability of bank failures over the next time period follows a logistic distribution that is estimated as follows:

$$P(Y_{it}=1)=\frac{1}{1+\exp(-\alpha-\beta X_{i,t-1})}, \tag{1}$$

where χ_{it} is an indicator variable that equals 1 if the bank is failed in time t, and $\chi_{i,t-1}$ is a vector of explanatory variables known at the end of the previous (or any appropriate lagged) period.

4.2 | Variable selection method

Although previous studies have introduced numerous variables to enhance the prediction accuracy of bank failure models, however which variables should be selected to predict failures in relative terms is inconclusive. The choice of variables is often driven by the popularity and/or significance of certain indicators across the literature. However, this is associated with the high risk of omitting unsuccessful variables in the past, which could be influential when confronted with new data. Thus, the selection of variables is useful to identify relevant variables and to enhance predictability (Tian, Yu, & Guo, 2015). Stepwise selection is a commonly used traditional variable selection approach that allows changes in either direction, dropping or adding one variable at a time according to some test statistics (Tian et al., 2015). However, it has a potential drawback. It ignores stochastic errors in the variable selection process (Fan & Li, 2001).

Consequently, we rely on univariate regression analysis for the selection of variables from a comprehensive list of 61 variables considered in the literature (see Table A1). Following Gupta et al. (2018), we perform univariate regression analysis of each of the 61 variables in turn, using the failure definition and econometric specification discussed earlier. To gauge the intertemporal discriminatory power of respective covariates we report regression estimates for 1-year (*T – 1*), 2-years (*T – 2*), and 3-years (*T – 3*) lagged time periods. To narrow down this list for further multivariate analysis, we exclude variables that (a) are not significant in all three time periods (to ensure that the selected covariates are consistent predictors of banks' financial soundness over a sufficiently long-time interval to allow for developing a reasonable early warning system), and (b) exhibit AMEs of less than 5% in t–1 time period. The rationale is that a unit change in the value of significant variables must induce sufficient change in the magnitude of the outcome probability to clearly distinguish between failed and non-failed banks (Gupta et al., 2018).

Most considered variables are statistically significant at the 1, 5, or 10% significance levels across all lagged time periods (see Table 2). However, only 19 out of 61 variables have AME values of 5% or more in t−1 time period. This suggests that, although these variables are significant predictors, a unit change in their value does not transmit significant change in the probability of outcome variable. Table 2 reports the final list of explanatory variables that we use for further multivariate regression analysis among all banks. An interesting observation in Table 2 is that the variable with the highest AME, net charge off to total assets (NCOTA), is largely ignored in the literature. Furthermore, the aggregated non-performing loans to total assets (NPLTA) ratio, which is considered to be one of the most common default predictors in the literature, has lower AME than one of its components (PD90TA) for the 1- and 2-years lagged periods, but higher AME for the 3-years lagged period. This indicates that the aggregated non-performing loans to total assets (NPLTA) is a superior predictor for bank failure in the longer time horizon (3 years and above).

We rerun the univariate regression analysis of each variable (total 61 variables) to verify its power to explain the failure of small, medium, and large banks respectively. Specifically, we verify whether the statistical significance of variables vary across size categories or not. Most of the considered variables are statistically significant at the 1, 5, or 10% significance levels in explaining the failure of small, medium, and large sized banks (see Table 3). Subsequently, we repeat the elimination process performed above using different bank size classifications (small, medium, and large). We find relatively similar results of univariate regression analysis compared with all banks, but different AME and their ranking, as well as additional variables across size categories. The final lists contain 19, 20, and 21 variables for small, medium, and large banks respectively. All of the 19 variables that we report as significant and that have AME of 5% or more in t−1time period for all banks (see Table 3) are the same across size categories, except the ratio total deposits to total assets (TDTA), which is rejected among large banks. Furthermore, we find additional variables such as NIM that meet the criteria for medium and large banks. Table 3 reports the final list of variables that we use for further multivariate regression analysis for small, medium, and large banks.

A noteworthy observation in Table 3 is that the AMEs of small banks' variables are mostly the highest for the 1-year lagged estimate. However, the ranking is changed for the second- and third-year lagged periods. The variables of large banks have the highest AMEs. This implies that the variables of small banks tend to have a strong prediction on a shorter horizon, while the variables of large banks tend to have a stronger prediction power in the longer horizon. Overall, these findings strongly support our belief that the magnitudes (AMEs) of mutually significant factors explaining bank failures vary across small, medium, and large size categories.

4.3 | Multivariate model-building strategy

Although several studies attempted to develop a parsimonious model that is numerically stable and applicable, they lack consensus on the criteria for including a variable in the multivariate model. According to Hosmer, Lemeshow, and Sturdivant (2013), the *SE* of a multivariate regression model increases with the number of variables and makes the model more dependent on the observed data. Thus, we use the approach suggested by Gupta et al. (2018) to minimize the number of explanatory variables entering the multivariate models. This approach requires the ranking of variables in univariate regression (reported in Tables 2 and 3) based on the magnitude of their AME (the variable with the highest value of AME for 1-year lagged (t −1) is ranked 1, and so on), and then each variable is introduced in turn into the multivariate model in declining order of their respective AME. Gupta et al. (2018) justify that the higher the value of AME, the higher the change in the predicted probability due to unit changes in the variable's value. In addition, a variable with a higher value of AME (e.g., NCOTA in Table 2) is more efficient than a variable with a lower value of AME (e.g., NIETA in Table 2) in discriminating between failed and non-failed banks. Thus, a covariate with a higher AME should have a priority entry in the multivariate prediction model. We also exclude a variable from the multivariate models if, when added it (a) changes the sign of any previously added variable, (b) holds the opposite sign to that generated by univariate analysis, (c) holds the identical sign to univariate analysis, but is insignificant with a p-value greater than 0.10, and (d) makes a previously introduced variable insignificant with a p-value greater than 0.10. This screening mechanism ensures that the method is useful to appropriately address the issue of multicollinearity, and gives a parsimonious multivariate model. Using panel logistic regression, this process is applied to all, small, medium, and large banks respectively for all three ($T - 1$, $T - 2$, and $T - 3$) respective lagged time periods. We do this to observe any variances that may arise due to different estimation models across different size classes.

We eventually end up with six variables to be used in the multivariate model for all banks. The variables are net charge off (NCOTA), past due 90+ days (PD90TA), loan loss reserves (LLRTA), total equity (TETA), other real estate owned (OREOTA), and total of non-interest expense (NIETA), and they are expressed as a ratio with respect to the bank's total assets. For small banks, the multivariate regression model is explained by seven variables. Five out of the seven variables (NCOTA, PD90TA, LLRTA, OREOTA, and NIETA) are common to explanatory variables for all banks. The other two variables are total deposits to total assets ratio (TDTA) and total interest expenses to total liabilities (TIETLB). Among large banks, five variables (PD90TA, LLRTA, TETA, NIM, and Loan Loss Provisions to Total Loans, LLPTL) are included in the multivariate regression model. Only three variables (PD90TA, LLRTA, and TETA) are similar to the variables of all banks. For medium banks, the multivariate regression model contains six variables as a combination of the explanatory variables for small and large banks (NCOTA, PD90TA, LLRTA, TETA, OREOTA, and NIM).

To further evaluate the consistency and strength of respective sets of main variables in jointly predicting the probability of banks' failures, we estimate another set of multivariate models supplementing control variables (discussed in Section 3.4.3). This helps us to control for potential differences in bank stability, banking crises, and state-level economic conditions. To gauge their intertemporal predictive ability, we also estimate regression models for 2- and 3-years lagged periods. The models and their results are presented in Tables 4–6.

4.4 | Multivariate regression results and discussion

4.4.1 | All banks

The results in columns 2, 3 and 4 of Table 4 indicate that the coefficients on NCOTA, PD90TA, LLRTA, and OREOTA have a positive influence on the probability of failure, implying that a weaker asset quality is associated with a higher bank failure. This is consistent with Imbierowicz and Rauch (2014), who find that credit risk has a prominent role in the overall stability of a bank. The coefficient on NIETA is positively related to bank failures. This suggests that a high level of bank operating expenses increases the likelihood of failure. This is in line with the findings of DeYoung (1998) who shows that poor management reduces the efficiency of using resources, thereby increasing the probability of default. In contrast, the coefficient on TETA is negative, suggesting that a higher capital is associated with a lower probability of failure. This is intuitive as the capital serves as a main line of defence against bank failures (Berger & Bouwman, 2013). All of these results are supported by several studies within the theoretical literature (e.g., Bryant, 1980; Repullo, 2004).

Turning to the control variables, house price inflation shows significantly negative values for all three lagged periods. This implies that declining real estate prices increase the probability of bank failures. This result is similar to the findings of Berger et al. (2016) who report that house price inflation has a negative effect, mostly on the 2 years preceding the failure. In contrast, foreign ownership is positively related to bank failures, suggesting that banks are more likely to fail if they have a greater percentage of foreign ownership. This result is in line with the findings of Berger et al. (2000) who show that foreign banks are generally less efficient than domestic banks in the US. The banking crises (SL and GFC) and primary regulator (FED and FDIC) dummies have significant and positive values for all lagged periods. Overall, the baseline model is parsimonious and offers a good model that fits the data. This is illustrated by, for example, the McKelvey and Zavoina pseudo R-squared, which is 77%. This value outperforms similar models in the early warning system literature (e.g., Cole & White, 2012; Poghosyan & Čihak, 2011). Additionally, the results of area under ROC (AUROC) curves of multivariate model for all banks, as shown in Appendix A1, exhibit that our models are excellent (around or above 90%) in classifying within-sample bank failures across all lagged time periods. However, AUROC values of the hold-out sample vary across different forecast horizons. The lowest estimate is 73% for the 3 years prior to the forecasting horizon, which is considered to be acceptable, while the 1- and 2-year forecast horizons are above 91%, suggesting excellent classification performance of our multivariate models.

4.4.2 | Small banks

Table 5 (columns 3, 7 and 11) reports the results of the main variables of multivariate regression models for small banks. NCOTA, PD90TA, LLRTA, OREOTA, and NIETA are identical to the multivariate regression models for all banks, are statistically significant, and have signs consistent with univariate regression estimates reported earlier. The other two variables are total deposits to total assets ratio (TDTA) as a measure of funding, and total interest expenses to total liabilities (TIETLB) as a proxy of liquidity. The coefficient on TDTA is significantly positive, suggesting that higher

TABLE 2 Univariate regression analysis of all banks

Variable (1)	1 year lag (2)	2 years lag (3)	3 years lag (4)	Rank (5)
TETA				5
β	−118.0046[a]	−78.3307[a]	−26.8250[a]	
SE	1.958	1.991	1.437	
AME%	−52.21[a]	−12.44[a]	−1.40[a]	
NPACR				13
β	−61.8227[a]	−59.8196[a]	−32.0309[a]	
SE	1.4194	1.5158	0.9211	
AME%	−23.19[a]	−10.53[a]	−1.92[a]	
LLRTA				4
β	226.1444[a]	202.4610[a]	132.0195[a]	
SE	4.5209	4.2604	4.7299	
AME%	58.06[a]	21.11[a]	2.60[a]	
PD90TA				2
β	132.8156[a]	128.1063[a]	107.1189[a]	
SE	2.5497	3.4981	4.8608	
AME%	68.22[a]	19.31[a]	2.35[a]	
NAATA				9
β	93.0304[a]	95.5749[a]	64.8830[a]	
SE	1.7678	2.0556	1.8676	
AME%	35.26[a]	9.25[a]	4.31[a]	
OREOTA				12
β	133.7940[a]	130.1982[a]	77.7526[a]	
SE	3.0969	2.7788	2.4406	
AME%	28.78[a]	8.81[a]	5.08[a]	
NPATA				15
β	61.4833[a]	68.4851[a]	43.4345[a]	
SE	1.3492	1.6684	1.0674	
AME%	21.43[a]	9.11[a]	2.92[a]	
LLPTL				14
β	58.7813[a]	54.5288[a]	36.4989[a]	
SE	0.9237	1.2300	1.3511	
AME%	22.87[a]	7.97[a]	2.96[a]	
LLPTA				3
β	120.3485[a]	115.2256[a]	93.5506[a]	
SE	1.5309	2.2250	2.9432	
AME%	62.07[a]	29.44[a]	2.67[a]	
NPLTL				17
β	44.6085[a]	41.7751[a]	26.9399[a]	
SE	0.8300	0.9684	0.8418	
AME%	14.73[a]	4.82[a]	1.95[a]	
NPLTA				8
β	77.9467[a]	79.7153[a]	59.6845[a]	

(Continues)

TABLE 2 (Continued)

Variable (1)	1 year lag (2)	2 years lag (3)	3 years lag (4)	Rank (5)
SE	1.2544	1.7377	1.6073	
AME%	35.36[a]	14.33[a]	5.01[a]	
NCOTA				1
β	142.3555[a]	130.1408[a]	91.1873[a]	
SE	1.9951	2.8493	3.0657	
AME%	68.69[a]	20.40[a]	7.06[a]	
NCOTL				10
β	78.7936[a]	69.5501[a]	45.6649[a]	
SE	1.2693	1.7141	1.7668	
AME%	34.64[a]	8.03[a]	3.22[a]	
ROA				7
β	−95.7321[a]	−76.4859[a]	−58.5776[a]	
SE	1.1160	1.5979	1.8411	
AME%	−48.40[a]	−32.34[a]	−3.70[a]	
TIETLB				16
β	50.5999[a]	45.5311[a]	33.0639[a]	
SE	1.6673	1.8924	2.2087	
AME%	15.80[a]	5.70[a]	0.70[a]	
TDTA				11
β	52.3850[a]	15.0219[a]	3.7117[a]	
SE	1.0468	0.7647	0.5178	
AME%	30.77[a]	1.51[a]	0.20[a]	
TLBTA				6
β	117.2084[a]	77.9171[a]	26.8385[a]	
SE	1.9600	2.0430	1.4386	
AME%	51.65[a]	12.67[a]	1.42[a]	
DIR				18
β	47.7084[a]	44.6106[a]	34.7815[a]	
SE	1.5661	1.8040	2.1400	
AME%	14.31[a]	5.42[a]	0.76[a]	
NIETA				19
β	78.0376[a]	53.8628[a]	32.1676[a]	
SE	2.0561	2.0986	2.1687	
AME%	13.32[a]	3.70[a]	2.51[a]	

Notes: a (b) [c] significant at the 1% (5%) [10%] level (two-sided test). The table reports univariate panel logistic regression results of final set of variables that we use for multivariate logit regression analysis. This excludes variables that are not significant in all three time periods or are significant but exhibit average marginal effects (AME) of less than 5% in all three time periods. β is the regression coefficient, *SE* is standard error and AME is average marginal effects in percentage. Ranking is based on the absolute values of AME for the 1-year lagged time estimate, where the highest value gets 1, second highest get 2 and so on.

deposits are associated with a higher probability of failure. This is consistent with Acharya and Naqvi (2012) who theoretically show that banks with excessive deposits are more likely to take risks by mitigating the lending standards to increase loans, because managers compensations are based on the volume of loans. It is

TABLE 3 Univariate regression analysis by size categories

Variable (1)	1 year lag: Small banks (2)	1 year lag: Medium banks (3)	1 year lag: Large banks (4)	2 years lag: Small banks (5)	2 years lag: Medium banks (6)	2 years lag: Large banks (7)	3 years lag: Small banks (8)	3 years lag: Medium banks (9)	3 years lag: Large banks (10)	Ranking: SB (11)	Ranking: MB (12)	Ranking: LB (13)
TETA										6	5	1
β	−104.5323[a]	−130.1084[a]	−120.8708[a]	−79.3857[a]	−100.3526[a]	−62.2437[a]	−27.4832[a]	−35.8680[a]	−22.1070[a]			
SE	3.2361	4.2753	3.1002	3.3538	3.6284	2.9461	2.4525	2.4180	2.9787			
AME%	−62.18[a]	−44.10[a]	−54.59[a]	−10.96[a]	−4.08[a]	−34.83[a]	−0.45[a]	−1.11[a]	−3.94[a]			
NPACR										13	12	16
β	−57.2967[a]	−68.9085[a]	−53.1749[a]	−54.6331[a]	−72.6386[a]	−46.1104[a]	−30.8899[a]	−47.1803[a]	−28.4403[a]			
SE	2.2476	2.9380	2.0797	2.1369	2.1867	2.1660	1.6521	2.0623	1.7933			
AME%	−29.20[a]	−21.26[a]	−20.61[a]	−13.27[a]	−7.71[a]	−14.17[a]	−1.06[a]	−0.15[a]	−6.58[a]			
LLRTA										5	4	5
β	213.0666[a]	240.7152[a]	231.7835[a]	198.8182[a]	286.6373[a]	176.7561[a]	141.3965[a]	173.0129[a]	78.7335[a]			
SE	8.0482	7.3646	10.1898	8.0616	8.4806	10.0358	8.7002	9.7912	10.1420			
AME%	63.01[a]	48.09[a]	47.76[a]	22.69[a]	6.29[a]	15.84[a]	4.50[a]	0.50[a]	4.56[a]			
PD90TA										10	3	3
β	133.9768[a]	139.6740[a]	117.3466[a]	128.8366[a]	140.7921[a]	104.5770[a]	113.9995[a]	117.5897[a]	92.9058[a]			
SE	4.9223	4.1417	6.7189	6.2654	5.6501	8.0909	7.2826	7.7970	10.1162			
AME%	43.65[a]	50.76[a]	51.97[a]	18.07[a]	11.99[a]	29.47[a]	7.59[a]	1.21[a]	12.05[a]			
NAATA										11	10	9
β	87.7152[a]	98.6135[a]	89.9165[a]	91.2369[a]	106.0918[a]	90.1672[a]	60.6484[a]	85.2727[a]	65.1436[a]			
SE	2.8367	3.0422	3.5646	3.5265	3.0938	4.1568	3.5138	3.7726	4.1951			
AME%	43.59[a]	28.57[a]	36.72[a]	6.58[a]	4.89[a]	15.74[a]	2.88[a]	0.50[a]	4.66[a]			
OREOTA										12	13	11
β	122.6557[a]	148.6665[a]	128.2099[a]	108.5423[a]	148.4729[a]	100.4943[a]	88.4473[a]	100.6918[a]	64.8086[a]			
SE	5.8518	4.2552	6.1331	4.1746	4.2838	5.5546	4.4715	5.2020	5.7444			
AME%	28.78[a]	20.01 [a]	27.70[a]	14.76[a]	4.83[a]	15.47[a]	4.21[a]	0.69[a]	5.60[a]			
NPATA										14	15	18
β	57.5061[a]	63.4320[a]	59.719[a]	67.9820[a]	73.0869[a]	58.4588[a]	45.2425[a]	58.9372[a]	40.8855[a]			
SE	2.0917	2.2385	3.1924	2.5115	2.0663	2.7041	2.0632	2.2254	2.4771			
AME%	27.42[a]	18.06[a]	17.85[a]	10.14[a]	5.97[a]	9.27[a]	3.03[a]	0.50[a]	2.58[a]			

(Continues)

TABLE 3 (Continued)

	1 year lag			2 years lag			3 years lag			Ranking		
Variable (1)	Small banks (2)	Medium banks (3)	Large banks (4)	Small banks (5)	Medium banks (6)	Large banks (7)	Small banks (8)	Medium banks (9)	Large banks (10)	SB (11)	MB (12)	LB (13)
LLPTL										15	14	15
β	50.3822[a]	62.6381[a]	71.1554[a]	52.4631[a]	59.8731[a]	58.3550[a]	36.1935[a]	43.3581[a]	32.1781[a]			
SE	1.4754	1.5675	2.5860	2.1884	1.9814	2.8988	2.2964	2.5763	3.2276			
AME%	24.71[a]	18.93[a]	21.14[a]	3.94[a]	4.69[a]	10.59[a]	1.87[a]	0.44[a]	3.16[a]			
LLPTA										2	2	4
β	111.6439[a]	125.6934[a]	130.3305[a]	112.8078[a]	123.4805[a]	117.4030[a]	92.0200[a]	93.5506[a]	93.5506[a]			
SE	2.3703	2.1606	4.2779	3.9636	3.7424	5.2697	4.5473	2.9432	2.9432			
AME%	81.30[a]	56.30[a]	48.55[a]	29.72[a]	21.26[a]	24.12[a]	7.86[a]	2.67[a]	2.67[a]			
NPLTL										17	16	19
β	37.7979[a]	48.1177[a]	52.6437[a]	33.6181[a]	47.8420[a]	43.2023[a]	24.2412[a]	36.2740[a]	27.6767[a]			
SE	1.2911	1.4320	2.3153	1.3998	1.4320	2.0705	1.4766	1.7402	1.9735			
AME%	14.77[a]	11.87[a]	14.32[a]	5.50[a]	2.50[a]	7.06[a]	1.19[a]	0.22[a]	2.17[a]			
NPLTA										9	9	10
β	72.1163[a]	81.9581[a]	80.1566[a]	76.6422[a]	88.3124[a]	77.8734[a]	59.0820[a]	78.4115[a]	57.4972[a]			
SE	2.0370	2.1214	3.1803	3.2444	2.7119	3.5575	2.8788	3.1037	3.5359			
AME%	46.56[a]	30.02[a]	31.61[a]	15.90[a]	8.81[a]	15.62[a]	4.63[a]	0.75[a]	5.26[a]			
NCOTA										1	1	6
β	134.5371[a]	145.7577[a]	156.7221[a]	128.4773[a]	143.0513[a]	126.1068[a]	97.0111[a]	108.7911[a]	75.2839[a]			
SE	2.8553	3.2351	5.8626	4.9892	4.5495	6.4348	5.3360	5.7823	7.4377			
AME%	96.88[a]	63.92[a]	47.15[a]	21.35[a]	11.32[a]	20.64[a]	5.72[a]	0.93[a]	6.36[a]			
NCOTL										8	11	12
β	71.0714[a]	84.7499[a]	91.1711[a]	68.6089[a]	77.1802[a]	69.2162[a]	47.4138[a]	53.2410[a]	39.9651[a]			
SE	1.5848	2.3689	3.4579	2.8728	2.6383	3.6113	2.9994	3.3405	4.3282			
AME%	53.11[a]	24.72[a]	26.30[a]	5.60[a]	3.43[a]	12.30[a]	2.38[a]	0.35[a]	3.70[a]			
ROA										4	7	7
β	−86.9303[a]	−102.0375[a]	−103.9933[a]	−82.1382[a]	−78.7770[a]	−91.1763[a]	−76.9376[a]	−56.2362[a]	−53.4303[a]			
SE	1.9747	1.7691	2.3017	3.2715	2.5243	4.1054	3.7828	2.8210	4.2183			
AME%	−63.23[a]	−43.17[a]	−43.95[a]	−26.60[a]	−23.25[a]	−22.14[a]	−2.62[a]	−2.10[a]	−7.34[a]			

TABLE 3 (Continued)

Variable (1)	1 year lag: Small banks (2)	1 year lag: Medium banks (3)	1 year lag: Large banks (4)	2 years lag: Small banks (5)	2 years lag: Medium banks (6)	2 years lag: Large banks (7)	3 years lag: Small banks (8)	3 years lag: Medium banks (9)	3 years lag: Large banks (10)	Ranking: SB (11)	Ranking: MB (12)	Ranking: LB (13)
TIETLB										18	19	14
β	61.4023[a]	47.7399[a]	40.9523[a]	57.8087[a]	44.3314[a]	39.8235[a]	36.4754[a]	36.8810[a]	41.0876[a]			
SE	3.1827	2.5850	3.1309	3.6377	2.9744	3.3368	4.0068	3.4618	3.7925			
AME%	14.07[a]	6.76[a]	22.92[a]	6.90[a]	1.98[a]	14.84[a]	1.70[a]	0.37[a]	7.04[a]			
TDTA										3	8	
β	93.2730[a]	73.8422[a]		46.6249[a]	24.8628[a]		14.1678[a]	6.6342[a]				
SE	2.6247	1.8049		2.4809	1.5792		1.7131	1.0472				
AME%	65.22[a]	35.73[a]		3.34[a]	0.69[a]		0.66[a]	0.03[a]				
TLBTA										7	6	2
β	104.0208[a]	129.7856[a]	118.6901[a]	79.5195[a]	91.3966[a]	61.8247[a]	26.9838[a]	35.8971[a]	22.1775[a]			
SE	3.2385	4.3062	3.0341	3.2817	2.9565	2.9201	2.4658	2.4175	2.9764			
AME%	61.66[a]	43.66[a]	53.71[a]	9.85[a]	8.57[a]	34.59[a]	0.66[a]	1.10[a]	3.99[a]			
DIR										19	18	17
β	57.6155[a]	46.8472[a]	36.4952[a]	55.6349[a]	44.4691[a]	37.9377[a]	33.6129[a]	38.6343[a]	43.0988[a]			
SE	2.9947	2.4835	2.7541	3.4871	2.8618	3.0714	3.8476	3.3762	3.6524			
AME%	12.90[a]	7.09[a]	20.42[a]	6.59[a]	2.26[a]	13.28[a]	1.43[a]	0.41[a]	6.87[a]			
NIETA										16	17	
β	93.6376[a]	92.4090[a]		75.5076[a]	52.9018[a]		63.1194[a]	26.8552[a]				
SE	4.1514	3.3347		3.5111	3.3639		3.8580	4.1414				
AME%	15.86[a]	7.23[a]		8.31[a]	1.79[a]		3.84[a]	0.24[a]				
NIM											20	8
β		−109.2913[a]	−108.9558[a]		−64.2987[a]	−72.3164[a]		−24.8567[a]	−34.0273[a]			
SE		8.0180	10.1674		7.3971	9.6534		7.4900	8.9746			
AME%		−6.11[a]	−38.72[a]		−2.79[a]	−24.13[a]		−0.75[a]	−9.27[a]			
CDLTA												20
β			10.2700[a]			12.6876[a]			14.1905[a]			
SE			0.5645			0.5577			0.5723			
AME%			5.96[a]			7.27[a]			8.00[a]			

(Continues)

TABLE 3 (Continued)

Variable (1)	1 year lag Small banks (2)	Medium banks (3)	Large banks (4)	2 years lag Small banks (5)	Medium banks (6)	Large banks (7)	3 years lag Small banks (8)	Medium banks (9)	Large banks (10)	Ranking SB (11)	MB (12)	LB (13)
TSTA												21
β			−10.3733[a]			−9.9772[a]			−10.3493[a]			
SE			0.5779			0.6435			0.7304			
AME%			−6.01[a]			−4.79[a]			−2.86[a]			
NIITA												13
β			−128.8201[a]			−90.1715[a]			−39.4006[a]			
SE			7.2345			7.4045			7.3900			
AME%			−24.09[a]			−12.03[a]			−3.10[a]			

Notes: a (b) [c] significant at the 1% (5%) [10%] level (two-sided test). The table reports univariate panel logistic regression results of the final set of variables for 1-year (columns 2 to 4), 2-years (columns 5 to 7), and 3-years (columns 8 to 10) lagged time periods across different size categories that we use for multivariate logit regression analysis. This excludes variables that are not significant in all three time periods or are significant but exhibit Average Marginal Effects (AME) of less than 5% in all three time periods. The sampling period is between 1985–2016. We consider banks corresponding to the bottom 25 percentile of total assets as small banks, those in the top 25 percentile as large banks, and the rest medium banks. If a bank fails in year t, the bank's failure indicator is "1" in that year t and "0" otherwise. β is the regression coefficient, *SE* is standard error and AME is Average Marginal Effects in percentage. Ranking (columns 11 to 13) is based on the absolute values of AME for the 1-year lagged time estimate for small banks (SB), medium banks (MB), and large banks (LB), where the highest value gets 1, second highest gets 2 and so on.

TABLE 4 Multivariate regression model for all banks

Panel A: Regression results						
	Without control variables			With control variables		
Variable (1)	1 year lag (2)	2 years lag (3)	3 years lag (4)	1 year lag (5)	2 years lag (6)	3 years lags (7)
NCOTA						
β	39.8244[a]	35.8465[a]	29.6089[a]	19.3895[a]	22.7565[a]	22.8679[a]
SE	2.6997	3.6711	4.1561	3.4203	4.1896	5.4609
AME%	15.0422[a]	6.1886[a]	2.5228[a]	5.3925[a]	5.1108[a]	2.2872[a]
PD90TA						
β	30.8333[a]	68.6989[a]	67.0855[a]	35.7708[a]	63.9088[a]	87.8005[a]
SE	3.6360	4.6782	4.7960	4.7021	5.5617	6.3381
AME%	11.6461[a]	11.8604[a]	5.7161[a]	9.9483[a]	14.3532[a]	8.7816[a]
LLRTA						
β	38.7496[a]	90.7056[a]	41.4054[a]	45.3049[a]	66.7501[a]	35.3956[a]
SE	3.8866	5.8187	5.9059	4.8571	6.4110	7.4069
AME%	14.6362[a]	15.6597[a]	3.5280[a]	12.5999[a]	14.9913[a]	3.5401[a]
TETA						
β	−75.4119[a]	−40.1401[a]	−12.6109[a]	−81.5346[a]	−48.3109[a]	−25.2473[a]
SE	2.2137	1.8575	1.2173	2.7717	2.4545	1.8338
AME%	−28.4841[a]	−6.9299[a]	−1.0745[a]	−22.6759[a]	−10.8501[a]	−2.5251[a]
OREOTA						
β	25.6350[a]	56.5833[a]	47.0656[a]	11.9870[a]	32.9740[a]	33.8065[a]
SE	1.9345	3.0603	2.8371	2.2337	3.3489	3.5696
AME%	9.6827[a]	9.7687[a]	4.0103[a]	3.3337[a]	7.4056[a]	3.3812[a]
NIETA						
β	3.4345[c]	3.5456	6.5127[b]	17.1175[a]	12.1606[a]	21.1530[a]
SE	1.8988	2.8338	2.8252	2.3013	2.8837	3.4003
AME%	1. 2972[c]	0.6121	0.5549[b]	4.7606[a]	2.7311[a]	2.1156[a]
GHPI						
β				−11.8082[a]	−12.2518[a]	−16.3320[a]
SE				0.7100	0.7338	0.8243
AME%				−3.2840[a]	−2.7516[a]	−1.6334[a]
SL						
β				2.3327[a]	2.7928[a]	2.2307[a]
SE				0.1377	0.1448	0.1479
AME%				0.6487[a]	0.6272[a]	0.2231[a]
GFC						
β				1.7208[a]	2.4787[a]	3.5303[a]
SE				0.1319	0.1392	0.1290
AME%				0.4786[a]	0.5566[a]	0.3531[a]
FOPCT						
β				2.7074[a]	3.1634[a]	3.2666[a]
SE				0.1368	0.14717	0.1289
AME%				0.7529[a]	0.7104[a]	0.3267[a]

(Continues)

TABLE 4 (Continued)

Panel A: Regression results						
	Without control variables			**With control variables**		
Variable (1)	**1 year lag (2)**	**2 years lag (3)**	**3 years lag (4)**	**1 year lag (5)**	**2 years lag (6)**	**3 years lags (7)**
FDIC						
β				2.8677[a]	2.4352[a]	1.4718[a]
SE				0.1399	0.1502	0.1426
AME%				0.7975[a]	0.5469[a]	0.1472[a]
FED						
β				2.8925[a]	2.4565[a]	1.4033[a]
SE				0.1853	0.1925	0.1974
AME%				0.8044[a]	0.5517[a]	0.1403[a]
Panel B: Goodness of fit measures						
Wald Chi2	2615[a]	1805[a]	1495[a]	1923[a]	1308[a]	1560[a]
Log likelihood	−4,722	−6,698	−7,189	−3,230	−4,606	−4,683
R2	0.7569	0.3018	0.0810	0.7722	0.5131	0. 2,998
No. of "0"	276,981	258,270	240,317	257,801	239,877	223,809
No. of "1"	1,694	1,554	1,342	1,546	1,337	1,040

Notes: a (b) [c] significant at the 1% (5%) [10%] level (two-sided test). Panel A presents multivariate panel logistic regression results for 1-year, 2-years, and 3-years lagged periods, estimated over a sampling period of 1985–2016. Columns 2, 3 and 4 do not include control variables and the rest include control variables in the multivariate estimates. If a bank fails in year t, the bank's binary indicator is "1" in that year t and "0" otherwise. A positive coefficient (β) suggests a positive relationship with failure likelihood and vice-versa. *SE* is standard error of respective coefficients and AME is Average Marginal Effects in percentage. Panel B reports the chi-square, the likelihood ratio and the coefficient of determination (McKelvey and Zavoina's R2) to measure the model's goodness of fit. No. of "1" counts the number of failures in our sample, while No. of "0" counts the number of "non-failure" observations.

also consistent with a recent empirical paper by Khan, Scheule, and Wu (2017) who find that banks holding higher deposits generally take more risks. This risk-taking can be attributed to the moral hazard of deposit insurance (Keeley, 1990). Moreover, we find the coefficient of TIETLB is significant and positively related to bank failures, implying that a higher share of interest expenses to total liabilities is associated with a higher probability of failure. This is in line with the findings of Betz et al. (2014) who show that the share of interest expenses to total liabilities has a positive effect on bank failures. These results are important to the literature in two ways. First, the low funding risk, as proxied by higher deposit ratios, has a more adverse effect on small banks and participates heavily in their failures. Second, the ratio of total interest expenses to total liabilities (TIETLB) contributes to explaining the relationship between liquidity risk and bank failures, specifically in small banks.

Next, we complement the models estimated in Table 5 with control variables (see Table 6). We find that all variables are statistically significant, and the sign of respective coefficients remains the same as the multivariate models estimated without control variables. An exception is NCOTA, which is insignificant for the 3-years lagged estimate. Furthermore, all control variables are statistically significant, and have a sign consistent with the control variables of the multivariate regression model for all banks.

The within-sample area under ROC (AUROC) curves of multivariate models developed across small banks are above 91%, suggesting excellent classification performance of our multivariate models for small banks across all time periods. The AUROC for out-of-sample for the 1- and 2-year horizons are excellent (above 83%), while that for the 3-year horizon is acceptable with 73% (see Appendix A1). These values and the shapes of ROC curves are relatively similar to the values and shapes of the ROC curves of all banks. This might indicate that small banks dominate the sample. Therefore, the effects of medium and large banks could be disregarded, thereby leading to a heterogeneous sampling and biased estimates. This

TABLE 5 Multivariate regression models without control variables by size categories

Panel A: Regression results												
	1 year lag				2 years lag				3 years lag			
Variable (1)	All banks (2)	Small banks (3)	Medium banks (4)	Large banks (5)	All banks (6)	Small banks (7)	Medium banks (8)	Large banks (9)	All banks (10)	Small banks (11)	Medium banks (12)	Large banks (13)
NCOTA												
β	39.8244[a]	47.7780[a]	26.7794[a]		35.8465[a]	35.2795[a]	24.3572[b]		29.6089[a]	34.9807[a]	18.6729[c]	
SE	2.6997	4.3729	6.9356		3.6711	5.8198	9.7469		4.1561	7.7074	9.7194	
AME%	15.0422[a]	25.9338[a]	8.2233[a]		6.1886[a]	9.2955[a]	2.4260[b]		2.5228[a]	2.3449[a]	0. 4449[c]	
PD90TA												
β	30.8333[a]	40.0387[a]	49.2210[a]	30.7840[b]	68.6989[a]	60.4810[a]	103.1499[a]	49.6830[a]	67.0855[a]	74.9498[a]	59.3297[a]	49.4819[a]
SE	3.6360	5.5941	9.7116	13.4527	4.6782	7.5487	12.4985	14.7570	4.7960	9.4391	12.5299	13.6310
AME%	11.6461[a]	21.7330[a]	15.1145[a]	14.3720[b]	11.8604[a]	15.9356[a]	10.2738[a]	17.2549[a]	5.7161[a]	5.0242[a]	1.4135[a]	21.8576[a]
LLRTA												
β	38.7496[a]	37.2199[a]	76.9432[a]	90.8796[a]	90.7056[a]	80.8907[a]	151.7489[a]	111.7549[a]	41.4054[a]	48.9785[a]	89.8852[a]	51.0923[a]
SE	3.8866	6.2611	11.4386	13.0387	5.8187	9.6021	14.3551	16.0917	5.9059	12.0932	14.1703	15.1547
AME%	14.6362[a]	20.2029[a]	23.6273[a]	42.4285[a]	15.6597[a]	21.3132[a]	15.1143[a]	38.8126[a]	3.5280[a]	3.2832[a]	2.1416[a]	22.5689[a]
TETA												
β	−75.4119[a]		−76.8456[a]	−58.5020[a]	−40.1401[a]		−46.1721[a]	−28.6889[a]	−12.6109[a]		−19.4217[a]	−16.7651[a]
SE	2.2137		7.2994	5.3171	1.8575		4.4201	3.5789	1.2173		2.9238	2.9133
AME%	−28.4841[a]		−23.5974[a]	−27.3126[a]	−6.9299[a]		−4.5988[a]	−9.9637[a]	−1.0745[a]		−0. 4627[a]	−7.4056[a]
OREOTA												
β	25.6350[a]	30.3981[a]	29.5583[a]		56.5833[a]	50.6729[a]	58.6815[a]		47.0656[a]	56.9582[a]	50.5365[a]	
SE	1.9345	2.9128	5.6556		3.0603	5.2876	7.1723		2.8371	6.1759	7.4430	
AME%	9.6827[a]	16.5000[a]	9.0766[a]		9.7687[a]	13.3513[a]	5.8447[a]		4.0103[a]	3.8181[a]	1.2040[a]	
NIETA												
β	3.4345[c]	24.7046[a]			3.5456	50.8636[a]			6.5127[b]	56.7940[a]		
SE	1.8988	3.2496			2.8338	5.1278			2.8252	5.8014		
AME%	1. 2972[c]	13.4096[a]			0.6121	13.4016[a]			0.5549[b]	3.8071[a]		
TDTA												
β		40.3383[a]				15.4303[a]				6.7069[a]		

(Continues)

TABLE 5 (Continued)

Panel A: Regression results												
	1 year lag				**2 years lag**				**3 years lag**			
Variable (1)	**All banks (2)**	**Small banks (3)**	**Medium banks (4)**	**Large banks (5)**	**All banks (6)**	**Small banks (7)**	**Medium banks (8)**	**Large banks (9)**	**All banks (10)**	**Small banks (11)**	**Medium banks (12)**	**Large banks (13)**
SE		2.6588				2.0183				1.4217		
AME%		21.8956[a]				4.0656[a]				0.4496[a]		
TIETLB												
β		6.4420[c]				25.9831[a]				10.8052[b]		
SE		3.3363				4.4912				5.2393		
AME%		3.4967[c]				6.8460[a]				0.7243[c]		
NIM												
β			−69.8740[a]	−61.2067[a]			−90.6152[a]	−77.2748[a]			−42.6528[a]	−37.2785[a]
SE			9.5580	9.3690			11.4190	10.8521			9.7505	8.6466
AME%			−21.4566[a]	−28.5753[a]			−9.0253[a]	−26.8376[a]			−1.0162[a]	−16.4670[a]
LLPTL												
β				14.8423[a]				18.7471[a]				6.1916
SE				4.0385				4.9228				5.4214
AME%				6.9293[a]				6.5109[a]				2.7350
Panel B: Goodness of fit measures												
Wald Chi2	2615[a]	1134[a]	184[a]	332[a]	1805[a]	473[a]	614[a]	199[a]	1495[a]	599[a]	336[a]	106[a]
Log likelihood	−4,722	−1,561	−931	−580	−6,698	−1920	−1,691	−1,153	−7,189	−1859	−2043	−1,413
R2	0.7569	0.7765	0.6956	0.6850	0.3018	0.2818	0.2508	0.2723	0.0810	0.1003	0.0796	0.1141
No. of "0"	276,981	67,482	53,792	26,753	258,270	62,347	48,553	24,137	240,317	58,013	43,297	21,532
No. of "1"	1,694	573	347	198	1,554	515	398	256	1,342	403	403	276

Notes: a (b) [c] significant at the 1% (5%) [10%] level (two-sided test). Panel A presents multivariate panel logistic regression results without control variables for 1-year (columns 2 to 5), 2-years (columns 6 to 9), and 3-years (columns 10 to 13) lagged periods across different size categories. The sampling period runs between 1985–2016. We consider banks corresponding to the bottom 25 percentile of total assets as small banks, those in the top 25 percentile as large banks, and the rest medium banks. If a bank fails in year t, the bank's failure indicator is "1" in that year t and "0" otherwise. A positive coefficient (β) suggests a positive relationship with failure likelihood and vice-versa. *SE* is standard error of respective coefficients and AME is Average Marginal Effects in percentage. Panel B reports the chi-square, the likelihood ratio and the coefficient of determination (McKelvey and Zavoina's R2) to measure the model's goodness of fit. No. of "1" counts the number of failures in our sample, while No. of "0" counts the number of "non-failure" observations.

TABLE 6 Multivariate regression models with control variables by size categories

Panel A: Regression results												
	1 year lag				2 years lag				3 years lag			
Variable (1)	All banks (2)	Small banks (3)	Medium banks (4)	Large banks (5)	All banks (6)	Small banks (7)	Medium banks (8)	Large banks (9)	All banks (10)	Small banks (11)	Medium banks (12)	Large banks (13)
NCOTA												
β	19.3895[a]	23.6037[a]	16.1267[b]		22.7565[a]	16.8937[a]	7.9462		22.8679[a]	9.4457	11.5266	
SE	3.4203	5.9249	7.8013		4.1896	6.2052	7.1901		5.4609	6.9608	8.1313	
AME%	5.3925[a]	8.3371[a]	5.0714[b]		5.1108[a]	6.3238[a]	4.8672		2.2872[a]	3.6989	8.1205	
PD90TA												
β	35.7708[a]	36.8391[a]	29.7863[a]	74.9467[b]	63.9088[a]	41.6546[a]	41.1333[a]	23.3172	87.8005[a]	49.0316[a]	50.1830[a]	25.9338
SE	4.7021	7.3648	11.5435	18.1738	5.5617	7.8294	9.5826	20.0476	6.3381	8.2395	10.4303	19.1263
AME%	9.9483[a]	13.0121[a]	9.3670[a]	30.3851[a]	14.3532[a]	15.5926[a]	25.1950[a]	9.5130	8.7816[a]	19.2007[a]	35.3538[a]	24.0660
LLRTA												
β	45.3049[a]	55.4850[a]	74.4268[a]	48.0865[a]	66.7501[a]	63.1352[a]	75.4361[a]	109.3064[a]	35.3956[a]	32.9799[a]	52.4253[a]	58.7668[a]
SE	4.8571	7.9362	11.1310	14.5004	6.4110	8.6515	9.3138	21.5749	7.4069	9.1973	10.1822	14.3517
AME%	12.5999[a]	19.5980[a]	23.4054[a]	19.4953[a]	14.9913[a]	23.6334[a]	46.2061[a]	44.5949[a]	3.5401[a]	12.9149[a]	36.9335[a]	54.5342[a]
TETA												
β	−81.5346[a]		−80.0039[a]	−78.6671[a]	−48.3109[a]		−28.9044[a]	−32.2108[a]	−25.2473[a]		−13.3598[a]	−12.8896[a]
SE	2.7717		6.4577	5.8419	2.4545		2.9681	5.6366	1.8338		2.4391	3.1857
AME%	−22.6759[a]		−25.1593[a]	−31.8934[a]	−10.8501[a]		−17.7045[a]	−13.1414[a]	−2.5251[a]		−9.4119[a]	−11.9612[a]
OREOTA												
β	11.9870[a]	21.5538[a]	23.7403[a]		32.9740[a]	22.1463[a]	25.7580[a]		33.8065[a]	13.5405[a]	24.2017[a]	
SE	2.2337	3.5277	4.8817		3.3489	3.9491	4.1493		3.5696	4.3868	4.9287	
AME%	3.3337[a]	7.6131[a]	7.4657[a]		7.4056[a]	8.2900[a]	15.7773[a]		3.3812[a]	5.3024[a]	17.0500[a]	
NIETA												
β	17.1175[a]	41.0303[a]			12.1606[a]	44.0242[a]			21.1530[a]	43.3831[a]		
SE	2.3013	3.8786			2.8837	4.6609			3.4003	4.8454		
AME%	4.7606[a]	14.4924[a]			2.7311[a]	16.4796[a]			2.1156[a]	16.9887[a]		
TDTA												
β		35.1497[a]				12.6272[a]				5.0683[a]		
SE		2.9340				1.8330				1.3056		

(Continues)

TABLE 6 (Continued)

Panel A: Regression results												
	1 year lag				2 years lag				3 years lag			
Variable (1)	All banks (2)	Small banks (3)	Medium banks (4)	Large banks (5)	All banks (6)	Small banks (7)	Medium banks (8)	Large banks (9)	All banks (10)	Small banks (11)	Medium banks (12)	Large banks (13)
AME%		12.4153[a]				4.7267[a]				1.9847[a]		
TIETLB												
β		56.1849[a]				61.8262[a]				71.8786[a]		
SE		5.5543				5.5448				5.6244		
AME%		19.8453[a]				23.1434[a]				28.1475[a]		
NIM												
β			−31.9666[a]	−8.4281			−56.3848[a]	−77.9860[a]			−35.9209[a]	−26.9491[a]
SE			9.7895	11.9581			7.9308	14.1350			7.9246	9.1266
AME%			−10.0527[a]	−3.4169			−34.5369[a]	−31.8168[a]			−25.3062[a]	−25.0081[a]
LLPTL												
β				18.0002[a]				20.5102[a]				20.4884[a]
SE				11.9581				6.1476				5.0810
AME%				7.2977[a]				8.3678[a]				19.0128[a]
GHPI												
β	−11.8082[a]	−8.0294[a]	−13.0414[a]	−10.1367[a]	−12.2518[a]	−14.9996[a]	−10.3922[b]	−9.2679[a]	−16.3320[a]	−16.8950[a]	−16.8401[a]	−12.8330[a]
SE	0.7100	1.6400	1.4633	1.3588	0.7338	1.4891	0.9543	1.4679	0.8243	1.4993	1.0103	1.2233
AME%	−3.2840[a]	−2.8361[a]	−4.1012[a]	−4.1096[a]	−2.7516[a]	−5.6148[a]	−6.3654[a]	−3.7811[a]	−1.6334[a]	−6.6160[a]	−11.8638[a]	−11.9087[a]
SL												
β	2.3327[a]	2.2009[a]	2.1595[a]	−0.9551[b]	2.7928[a]	2.4410[a]	4.5441[a]	3.9445[a]	2.2307[a]	2.3423[a]	4.2429[a]	3.5135[a]
SE	0.1377	0.2429	0.3850	0.4718	0.1448	0.2372	0.3478	0.4437	0.1479	0.2350	0.3504	0.3618
AME%	0.6487[a]	0.7773[a]	0.6791[a]	−0.3872[b]	0.6272[a]	0.9137[a]	2.7833[a]	1.6093[a]	0.2231[a]	0.9172[a]	2.9891[a]	3.2605[a]
GFC												
β	1.7208[a]	1.6672[a]	1.7609[a]	1.7093[a]	2.4787[a]	1.6290[a]	2.0769[a]	3.0810[a]	3.5303[a]	2.9303[a]	3.3281[a]	4.8514[a]
SE	0.1319	0.2904	0.2593	0.3044	0.1392	0.2998	0.2174	0.3424	0.1290	0.2600	0.2226	0.2969
AME%	0.4786[a]	0. 5888[a]	0.5537[a]	0.6930[a]	0.5566[a]	0.6098[a]	1.2722[a]	1.2570[a]	0.3531[a]	1.1475[a]	2.3447[a]	4.5020[a]
FOPCT												
β	2.7074[a]	3.8462[a]	2.1136[a]	2.0205[a]	3.1634[a]	4.4577[a]	1.8411[a]	2.6174[a]	3.2666[a]	4.6263[a]	2.0534[a]	3.3944[a]

TABLE 6 (Continued)

Panel A: Regression results												
	1 year lag				2 years lag				3 years lag			
Variable (1)	All banks (2)	Small banks (3)	Medium banks (4)	Large banks (5)	All banks (6)	Small banks (7)	Medium banks (8)	Large banks (9)	All banks (10)	Small banks (11)	Medium banks (12)	Large banks (13)
SE	0.1368	0.2568	0.3237	0.4021	0.14717	0.2738	0.2120	0.3828	0.1289	0.2598	0.2167	0.2461
AME%	0.7529[a]	1.3585[a]	0.6647[a]	0.8191[a]	0.7104[a]	1.6686[a]	1.1277[a]	1.0678[a]	0.3267[a]	1.8116[a]	1.4466[a]	3.1499 [a]
FDIC												
β	2.8677[a]	3.9195[a]	1.9182[a]	0.1997	2.4352[a]	3.0064[a]	2.9880[a]	1.3465[a]	1.4718[a]	1.8555[a]	2.0219[a]	0.3370
SE	0.1399	0.2334	0.3215	0.2826	0.1502	0.2303	0.3095	0.3201	0.1426	0.2292	0.2945	0.2151
AME%	0.7975[a]	1.3844[a]	0.6032[a]	0.0809	0.5469[a]	1.1254[a]	1.8302[a]	0.5493[a]	0.1472[a]	0.7266[a]	1.4244[a]	0.3127
FED												
β	2.8925[a]	3.8668[a]	1.9393[a]	0.4448	2.4565[a]	3.2687[a]	2.7800[a]	1.1506[a]	1.4033[a]	2.1849[a]	1.7908[a]	0.3549
SE	0.1853	0.3366	0.4273	0.3625	0.1925	0.3124	0.3721	0.7462	0.1974	0.3092	0.3535	0.2743
AME%	0.8044[a]	1.3658[a]	0.6098[a]	0.1803	0.5517[a]	1.2236[a]	1.7028[a]	0.4694[a]	0.1403[a]	0.8556[a]	1.2616[a]	0.3294
Panel B: Goodness of fit measures												
Wald Chi2	1923[a]	140[a]	364[a]	551[a]	1308[a]	655[a]	1025[a]	177[a]	1560[a]	617[a]	671[a]	421[a]
Log likelihood	−3,230	−952	−596	−392	−4,606	−1,064	−1,167	−843	−4,683	−1,101	−1,273	−803
R2	0.7722	0.7901	0.7872	0.7678	0.5131	0.6240	0.6088	0.4522	0. 2,998	0.5564	0.5107	0.5532
No. of "0"	257,801	62,329	45,696	22,627	239,877	57,999	40,643	20,122	223,809	54,301	36,202	18,030
No. of "1"	1,546	512	284	175	1,337	403	321	230	1,040	308	285	194

Panel C: Model performance				
	All banks	**Small banks**	**Medium banks**	**Large banks**
AUROC-W	0.9805	0.9767	0.9864	0.9709
AUROC-H	0.9785	0.9077	0.9212	0.9869

Notes: a (b) [c] significant at the 1% (5%) [10%] level (two-sided test). Panel A presents multivariate panel logistic regression results with control variables for 1-year (columns 2 to 5), 2-years (columns 6 to 9), and 3-years (columns 10 to 13) lagged periods across different size categories. The sampling period runs between 1985–2016. We consider banks corresponding to the bottom 25 percentile of total assets as small banks, those in the top 25 percentile as large banks, and the rest medium banks. If a bank fails in year t, the bank's failure indicator is "1" in that year t and "0" otherwise. A positive coefficient (β) suggests a positive relationship with failure likelihood and vice-versa. *SE* is standard error of respective coefficients and AME is Average Marginal Effects in percentage. Panel B reports the chi-square, the likelihood ratio and the coefficient of determination (McKelvey and Zavoina's R2) to measure the model's goodness of fit. No. of "1" counts the number of failures in our sample, while No. of "0" counts the number of "non-failure" observations. Panel C shows the accuracy of models' performance measured by area under the ROC curve. AUROC-W represents within sample and AUROC-H represents hold-out sample area under ROC curves.

TABLE 7 Multivariate regression models with interaction between bank size and bank charted

Panel A: Regression results																		
	Without control variables									With control variables								
	1 year lag			2 years lag			3 years lag			1 year lag			2 years lag			3 years lag		
Variable (1)	β (2)	SE (3)	AME% (4)	β (5)	SE (6)	AME % (7)	β (8)	SE (9)	AME % (10)	β (11)	SE (12)	AME% (13)	β (14)	SE (15)	AME % (16)	β (17)	SE (18)	AME% (19)
NCOTA	38.998[a]	2.7118	14.73[a]	35.666[a]	3.7124	5.99[a]	29.939[a]	4.1816	2.66[a]	27.606[a]	3.8695	6.50[a]	31.308[a]	4.5665	6.74[a]	33.890[a]	6.4200	2.55[a]
PD90TA	30.512[a]	3.6420	11.52[a]	67.238[a]	4.7684	11.30[a]	66.400[a]	4.8608	5.90[a]	41.919[a]	5.3508	9.87[a]	70.985[a]	5.8977	15.27[a]	100.822[a]	7.4880	7.58[a]
LLRTA	40.001[a]	3.9157	15.11[a]	91.097[a]	5.9487	15.31[a]	40.211[a]	5.9427	3.58[a]	37.022[a]	5.5305	8.71[a]	60.303[a]	6.6313	12.97[a]	30.747[a]	8.7848	2.31[a]
TETA	−75.874[a]	2.2259	−28.65[a]	−40.615[a]	1.8965	−6.83[a]	−12.508[a]	1.2240	−1.11[a]	−83.817[a]	2.9414	−19.73[a]	−46.240[a]	2.2630	−9.95[a]	−26.354[a]	2.0396	−1.98[a]
OREOTA	25.653[a]	1.9475	9.69[a]	58.195[a]	3.1593	9.78[a]	48.086[a]	2.8721	4.28[a]	16.834[a]	2.5731	3.96[a]	37.812[a]	3.3371	8.13[a]	42.443[a]	4.4424	3.19[a]
NIETA	3.459[c]	1.9709	1.31[c]	−0.480	2.9576	−0.08	2.792	2.8957	0.25	15.257[a]	2.6780	3.59[a]	7.812[a]	3.0852	1.68[a]	15.721[a]	4.0011	1.18[a]
MB	−0.276[b]	0.1330	−0.11[a]	−1.272[a]	0.1888	−0.15[a]	−1.264[a]	0.1686	−0.07[a]	−0.218	0.1692	−0.09[a]	−0.751[a]	0.1895	−0.09[a]	−1.395[a]	0.2405	−0.05[a]
LB	0.057	0.1457	−0.03	−1.108[a]	0.2077	−0.11[a]	−1.110[a]	0.1869	−0.05[a]	0.511[a]	0.1843	−0.02	−0.107	0.2002	−0.03	−1.022[a]	0.2650	−0.03[a]
SC	0.316[b]	0.1258	0.09[a]	−0.953[a]	0.1836	−0.08[a]	−1.487[a]	0.1684	−0.08[a]	−6.359[a]	0.8218	−10.75[a]	−7.894[a]	0.8515	−18.5[a]	−9.594[a]	1.1271	−16.2[a]
MB × SC	−0.040	0.1602		0.845[a]	0.2230		1.099[a]	0.2052		−0.255	0.2145		0.583[a]	0.2301		1.353[a]	0.2894	
LB × SC	−0.194	0.1789		0.947[a]	0.2512		1.259[a]	0.2315		−1.049[a]	0.2402		−0.012	0.2497		1.233[a]	0.3239	
GHPI										−11.935[a]	0.7770	−2.81[a]	−11.666[a]	0.7378	−2.51[a]	−17.606[a]	1.0134	−1.32[a]
FOPCT										2.673[a]	0.1496	0.63[a]	3.074[a]	0.1427	0.66[a]	3.449[a]	0.1621	0.26[a]
SL										2.595[a]	0.1441	0.61[a]	3.319[a]	0.1440	0.71[a]	2.991[a]	0.1619	0.22[a]
GFC										1.790[a]	0.1390	0.42[a]	2.576[a]	0.1403	0.55[a]	3.764[a]	0.1542	0.28[a]
FED										9.141[a]	0.8382	2.15[a]	9.295[a]	0.8519	2.00[a]	9.138[a]	1.1092	0.69[a]
FDIC										9.077[a]	0.8261	2.14[a]	9.222[a]	0.8411	1.98[a]	9.191[a]	1.0954	0.69[a]
Panel B: Goodness of fit measures																		
Wald Chi2			2616[a]			1692[a]			1540[a]			1586[a]			1387[a]			1187[a]
Log likelihood			−4,689			−6,665			−7,137			−2,941			−4,234			−4,408
R^2			0.7599			0.299			0.0889			0.8093			0.6433			0.4554
No. of “0”			276,973			258,269			240,317			257,800			239,877			223,809
No. of “1”			1,687			1,554			1,342			1,546			1,337			1,040

Notes: a (b) [c] significant at the 1% (5%) [10%] level (two-sided test). Panel A presents multivariate panel logistic regression results with interaction terms (between bank size and the bank charted) for 1-, 2-, and 3-years lagged periods. Size category “Small Banks” and bank charted “National Charted” are considered reference groups, and thus main and interaction effects are reported for medium banks (MB), large banks (LB) and State charted banks (SC). Results are reported separately for multivariate models without control variables (columns 2 to 10) and with control variables (columns 11 to 19). The sampling period runs between 1985–2016. We consider banks corresponding to the bottom 25 percentile of total assets as small banks, those in the top 25 percentile as large banks, and the rest medium banks. If a bank fails in year t, the bank's failure indicator is “1” in that year t and “0” otherwise. A positive coefficient (β) suggests a positive relationship with failure likelihood and vice-versa. *SE* is standard error of respective coefficients and AME is average marginal effects in percentage. Panel B reports the chi-square, the likelihood ratio and the coefficient of determination (McKelvey and Zavoina's R2) to measure the model's goodness of fit. No. of “1” counts the number of failures in our sample, while No. of “0” counts the number of “non-failure” observations.

TABLE 8 Multivariate regression models using alternative bank size cut-off

Panel A: Regression results						
	Without control variables			**With control variables**		
Variable (1)	**Small banks (2)**	**Medium banks (3)**	**Large banks (4)**	**Small banks (5)**	**Medium banks (6)**	**Large banks (7)**
NCOTA						
β	61.0199[a]	107.2142[a]		37.2589[a]	45.9734	
SE	2.7163	26.0847		3.2149	33.9338	
AME%	23.0172[a]	45.4571[a]		11.8165[a]	14.9293	
PD90TA						
β	34.2282[a]	76.9234[c]	0.0598	17.2857[a]	133.9925[a]	−3.3178
SE	3.6994	44.3149	43.2590	4.4076	54.1602	53.8908
AME%	12.9111[a]	32.6143[c]	0.0440	5.4821[a]	43.5124[a]	−3.2719
LLRTA						
β	79.9810[a]	33.3856	35.6659	79.9705[a]	108.3962[b]	−16.5536
SE	4.1900	35.7125	32.5191	4.3175	51.0881	44.9606
AME%	30.1695[a]	14.1549	26.2155	25.3623[a]	35.2003[b]	−16.3246
TETA						
β		−44.7655[a]	−23.2248[a]		−42.4832[a]	−22.3881[a]
SE		10.3977	7.6490		13.7155	8.7300
AME%		−18.9799[a]	−17.0709[a]		−13.7959[a]	−22.0784[a]
OREOTA						
β	42.1623[a]	0.2320		26.8148[a]	24.0758	
SE	2.0045	19.0124		1.9264	25.1644	
AME%	15.9039[a]	0.0984		8.5042[a]	7.8783	
NIETA						
β	16.4045[a]			24.4202[a]		
SE	2.0063			2.1098		
AME%	6.1879[a]			7.7447[a]		
TDTA						
β	17.7231[a]			18.6877[a]		
SE	1.0410			1.0974		
AME%	6.6853[a]			5.9467[a]		
TIETLB						
β	8.7079[a]			57.3154[a]		
SE	1.9568			2.9869		
AME%	3.2847[a]			18.1773[a]		
NIM						
β		−64.0055[b]	−30.3796		−56.3185	−30.3595
SE		28.0799	21.2082		39.2101	25.0292
AME%		−2.7137[b]	−22.3299		18.2887	−29.9395
LLPTL						
β			38.1682[a]			34.3973[b]
SE			11.1679			15.0606
AME%			28.0547[a]			33.9214[b]

(Continues)

TABLE 8 (Continued)

Panel A: Regression results						
	Without control variables			**With control variables**		
Variable (1)	**Small banks (2)**	**Medium banks (3)**	**Large banks (4)**	**Small banks (5)**	**Medium banks (6)**	**Large banks (7)**
GHPI						
Β				−11.4753[a]	−14.4901[a]	−10.2888[a]
SE				0.6820	3.2148	2.6440
AME%				−3.6393[a]	−4.7054[a]	−10.1465[a]
SL						
β				2.2653[a]		
SE				0.1438		
AME%				0.7184[a]		
GFC						
β				1.8375[a]	2.3089[a]	2.1645[b]
SE				0.1354	0.9283	1.0859
AME%				0.5827[a]	0.7497[a]	2.1346[b]
FOPCT						
β				3.8084[a]	2.4254[b]	
SE				0.1330	1.2692	
AME%				1.2078[a]	0.7876[b]	
FDIC						
β				3.4747[a]	−0.2916	−1.0419
SE				0.1349	0.7703	0.6427
AME%				1.1020[a]	0.0947	−1.0274
FED						
β				3.6334[a]	−0.2368	−1.6951
SE				0.1817	0.9574	1.0726
AME%				1.1523[a]	0.0769	−1.6716
Panel B: Goodness of fit measures						
Wald Chi2	2573[a]	127[a]	72[a]	4870[a]	71[a]	54[a]
Log likelihood	−5,270	−83	−106	−3,478	−52	−67
R2	0.4666	0.6345	0.3338	0.6849	0.6931	0. 4,963
No. of “0”	262,768	4,417	2,828	244,409	3,878	1,645
No. of “1”	1,620	31	24	1,478	28	21

Notes: a (b) [c] significant at the 1% (5%) [10%] level (two-sided test). Panel A presents multivariate panel logistic regression results with and without control variables for 1-year lagged periods across different size categories. The sampling period runs between 1985–2016. We consider small banks (total assets, or TA, up to $1 billion), medium banks (TA exceeding $1 billion and up to $3 billion), and large banks (TA exceeding $3 billion). If a bank fails in year t, the bank's failure indicator is “1” in that year t and “0” otherwise. A positive coefficient (β) suggests a positive relationship with failure likelihood and vice-versa. *SE* is standard error of respective coefficients and AME is Average Marginal Effects in percentage. Panel B reports the chi-square, the likelihood ratio and the coefficient of determination (McKelvey and Zavoina's R2) to measure the model's goodness of fit. No. of “1” counts the number of failures in our sample, while No. of “0” counts the number of “non-failure” observations.

clearly supports the necessity of distinction between different size classes when analysing bank failures.

4.4.3 | Medium banks

Table 5 (columns 4, 8 and 12) shows that five out of six main variables (NCOTA, PD90TA, LLRTA, OREOTA, and TETA) of the multivariate regression model for medium banks remain the same as the multivariate models estimated for all and small banks. They are statistically significant and have the expected sign of respective coefficients across all lagged periods. The sixth main variable is NIM, which is also statistically significant and has a negative sign across the three lagged periods, suggesting that a larger amount of returns generated by investments reduces the probability of failure for medium banks. This is consistent with the hypothesis that banks dealing heavily with risky loans tend to have higher net interest margins (Angbazo, 1997).

Table 6 (columns 4, 8 and 12) reports the results after introducing control variables. All variables (main and control) are statistically significant and all coefficients hold the same sign previously reported except NCOTA, which is insignificant for 2- and 3-years lagged estimates. Both within-sample and out-of-sample classification of all multivariate models across all time periods are above 81%, which is considered to be excellent (see Appendix A1).

4.4.4 | Large banks

As reported in Table 5 (columns 5, 9 and 13), multivariate regression models for large banks contain the ratio of loan loss provisions to total loans (LLPTL) as one of the main variables that has not been reported for all, small, and medium banks. The coefficient on LLPTL is positive and statistically significant for 1- and 2-years lagged estimates but becomes insignificant for the 3-years lagged estimate. This indicates that risky loan portfolios increase the probability of failure of large banks more than other banks. Similarly, Poghosyan and Čihak (2011) find that the deterioration of the loan portfolio enhances the probability of bank default. The rest of the main variables (PD90TA, LLRTA, TETA, and NIM) are statistically significant and have a sign consistent with univariate regression estimates across all three-time lagged periods.

In the presence of control variables, three out of the five main variables are statistically significant and have the same sign as those of large banks' multivariate models estimated without control variables across three lagged periods (see Table 6). However, of the other two variables, NIM is insignificant for the 1-year lagged period, and PD90TA is insignificant for 2- and 3-years lagged estimates. The control variables are statistically significant, and their coefficients have expected signs, except primary regulators (FED and FDIC) are insignificant for the 1- and 3-years lagged estimates.

The within-sample and out-of-sample AUROC estimated for multivariate models for large banks are close to, or higher than, 0.80, implying superior classification performance across all time periods (see Appendix A1). Yet the shapes of ROC curves of hold-out sample estimates are steps rather than concave, due to the scarcity of failures in out-of-sample validation.

5 | ROBUSTNESS CHECKS

5.1 | Interaction between bank size and bank charter

To test the hypothesis that the impact of bank size on the probability of bank failures varies with bank charter, we add interaction between bank size and bank charter to the multivariate regression models reported in Table 4. Table 7 reports the results of multivariate regression models with interaction terms for bank size and bank charter. These results are presented with and without control variables, and for the three lagged periods. The size category "Small Banks" and bank charter "National Chartered Banks" are taken as the reference group, and thus main and interaction effects are reported for medium banks, large banks and state-chartered banks. The notable result of interactions between bank size and bank charter is that all explanatory variables, as well as control variables, are statistically significant and have signs consistent with our expectation.[7] This shows the robustness and consistency of our explanatory variables.

The impact of medium sized banks (MB) is significantly negative across all estimates, but the main effect of large banks (LB) is only significantly negative for 2- and 3-years lagged estimates. These results are robust to the inclusion of control variables. The sign and statistically significant differences between medium and large banks for the 1-year lagged period, which is the main concern of this paper, confirms our main result that the probability of bank failures varies with size categories. The effects of state-chartered banks are significantly negative for all estimates with and without control variables.[8] This is mostly consistent with Danisewicz, McGowan, Onali, and Schaeck (2017), who show that the depositor preference law leads to less risk taking, and a lower probability of failure among state-chartered banks.[9]

TABLE 9 Financial crises and normal times

Panel A: Regression results																		
	All banks									Small banks								
	Banking crises			Market crises			Normal times			Banking crises			Market crises			Normal times		
Variable	β	SE	AME%	β	SE	AME%	β	SE	AME%	β	SE	AME%	β	SE	AME%	β	SE	AME%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
NCOTA	13.07[a]	4.58	9.28[a]	24.99[b]	11.54	3.76 [b]	28.22[a]	6.53	3.79[a]	20.71[b]	9.77	11.84[b]	23.68	17.48	6.51	26.43[a]	10.38	5.29[a]
PD90TA	17.34[a]	6.37	12.30[a]	28.49[b]	13.22	4.28[b]	39.96[a]	9.18	5.36[a]	33.32[a]	11.88	19.05[a]	45.38 [b]	19.36	12.47 [b]	10.52	14.20	2.11
LLRTA	48.11[a]	6.21	34.13[a]	69.45[a]	15.43	10.44[a]	54.85[a]	9.57	7.36[a]	55.07[a]	13.90	31.48[a]	40.08 [c]	22.40	11.01 [c]	89.77[a]	13.72	17.97[a]
TETA	−70.40[a]	3.11	−49.95[a]	−51.27[a]	7.66	−7.71[a]	−98.57[a]	5.19	−13.22[a]									
OREOTA	12.16[a]	2.89	8.63[a]	2.39	8.46	0.36	19.17[a]	4.06	2.57[a]	24.02[a]	5.91	13.73[a]	−0.80	13.16	−0.22	25.25[a]	5.68	5.06[a]
NIETA	7.84[a]	3.18	5.56[a]	36.45[a]	6.43	5.48[a]	15.91[a]	4.04	2.13[a]	33.84[a]	6.53	19.34[a]	45.72 [a]	9.82	12.56 [a]	41.70[a]	6.53	8.35[a]
TIETLB										47.45[a]	13.74	27.12[a]	17.64	21.22	4.85	28.48[a]	7.12	5.70[a]
TDTA										29.37[a]	4.87	16.79[a]	37.03 [a]	9.31	10.18 [a]	39.68[a]	4.75	7.94[a]
Panel B: Goodness of fit measures																		
Wald Chi2			2054[a]			518[a]			700[a]			442[a]			209 [a]			498[a]
Log likelihood			−1,685			−327			−907			−344			−140			−332
R^2			0.7776			0.8473			0.797			0.8458			0.8775			0.7709
No. of "0"			57,668			45,976			154,156			13,926			10,808			37,595
No. of "1"			908			218			420			269			115			128

Panel C: Regression results																		
	Medium banks									Large banks								
	Banking crises			Market crises			Normal times			Banking crises			Market crises			Normal times		
Variable (1)	β (2)	SE (3)	AME% (4)	β (5)	SE (6)	AME% (7)	β (8)	SE (9)	AME% (10)	β (11)	SE (12)	AME% (13)	β (14)	SE (15)	AME% (16)	β (17)	SE (18)	AME% (19)
NCOTA	9.96	13.12	5.17	−47.14	34.12	−9.43	33.05[a]	13.44	5.69[a]									
PD90TA	25.50	20.61	13.22	−10.10	26.88	−2.02	14.99	26.78	2.58	79.96[a]	26.04	96.53[a]	−960.30[b]	493.9	−9.78[b]	75.38	67.98	7.47
LLRTA	104.35[a]	20.49	54.12[a]	120.19[b]	51.23	24.05 [b]	56.08[a]	19.52	9.66[a]	28.76	18.75	34.73	430.60	291.87	4.39	147.42[c]	80.04	14.61[c]
TETA	−63.56[a]	9.46	−32.96[a]	−171.12[a]	48.29	−34.23 [a]	−102.02[a]	10.92	−17.58[a]	−59.96[a]	6.56	−72.39[a]	−625.78[b]	295.70	−6.37[b]	−123.09[b]	59.14	−12.20[b]
OREOTA	18.16[b]	8.43	9.42[b]	17.27	19.54	3.45	20.10[a]	8.19	3.46[a]									
NIM	−78.38[a]	17.85	−40.65[a]	−17.44	25.22	−3.49	−41.51[b]	20.30	−7.15[b]	−24.01[c]	14.92	−28.98[c]	543.82[c]	294.48	5.54[c]	−69.24	47.17	−6.86
LLPTL										28.52[a]	6.32	34.43[a]	41.54	81.17	0.42	−12.10	20.44	−1.20
Panel D: Goodness of fit measures																		
Wald Chi2			92[a]			17[b]			207[a]			294[a]			5			20[b]
Log likelihood			−241			−47			−174			−261			−13			−54
R^2			0.6818			0.9341			0.8411			0.6932			0.9213			0.8730
No. of "0"			10,191			6,847			28,656			5,054			3,467			14,103
No. of "1"			98			21			95			121			5			35

Notes: a (b) [c] significant at the 1% (5%) [10%] level (two-sided test). Panels A and C present the results of checks to establish the robustness of our results. The crises include banking crises (the credit crunch and the subprime lending crisis), market crises (the stock market crash; the Russian debt crisis; the dot.com bubble and September 11), and normal times. The sampling period runs between 1985–2016. We consider banks corresponding to the bottom 25 percentile of total assets as small banks, those in the top 25 percentile as large banks, and the rest medium banks. If a bank fails in year t, the bank's failure indicator is "1" in that year t and "0" otherwise. A positive coefficient (β) suggests a positive relationship with failure likelihood and vice-versa. *SE* is standard error of respective coefficients and AME is Average Marginal Effects in percentage. Panels B and D report the chi-square, the likelihood ratio and the coefficient of determination (McKelvey and Zavoina's R2) to measure the model's goodness of fit. No. of "1" counts the number of failures in our sample, while No. of "0" counts the number of "non-failure" observations.

Turning to the effects of bank size and bank charter, we observe a negative but insignificant relationship between medium sized banks and bank charter "MB × State Charter" for the 1-year lagged estimate. However, this relationship becomes positive and statistically significant for 2- and 3-years lagged estimates. For interaction terms between large sized banks and bank charter "LB × State Charter", we find relatively similar findings of "MB × State Charter". These results are robust to the presence of control variables. Overall, the impact of bank size on probability of bank failures varies with bank charter, and it might be appropriate to consider this when predicting the failure of US banks.

5.2 | Alternative size classification

One may argue to what extent the main results are driven by our definition of size classes. We rerun our analyses using Berger and Bouwman (2013) bank size cut-off, which is widely used in literature. They split the bank size classes into small banks (total assets, or TA, up to $1 billion), medium banks (TA exceeding $1 billion and up to $3 billion), and large banks (TA exceeding $3 billion). This reclassifies around 90% of our medium and large banks as small banks.

Table 8 shows the results using this alternative cut-off. Clearly, small banks results are similar to the main results. For medium banks, the results are relatively comparable to the main results. An obvious exception is the other real estate owned (OREOTA), which is insignificant with and without control variables. For large banks, only the capital (TETA) and the ratio of loan loss provisions to total loans (LLPTL) are similar to the main results. This inconsistency of results specifically for medium banks and more for large banks may attributed to the huge reduction in the bank failures sample impacted by different size cut-off.

5.3 | Crisis periods

According to Berger and Bouwman (2013), the effects of financial crises are likely to differ by crisis type. To test the reliability of our multivariate results, we examine bank failures during banking crises (the credit crunch and subprime lending crisis), market crises (the 1987 stock market crash, the 1998 Russian debt crisis and long-term capital management bailout, the dot.com bubble, and the September 11 terrorist attack [2000–2002]), and normal times (all non-crisis years) as three separate groups. We rerun all multivariate regressions separately for all, small, medium, and large banks with the same control variables used in the main multivariate regressions (see Section 3.4.3) with the exception of the credit crunch and subprime lending crisis, to avoid collinearity.

Table 9 reports the findings for all, small, medium, and large banks across various types of financial crises and normal times. For all banks, the results are the same at all times except that OREOTA becomes insignificant during the market crises. For small banks, all variables are significant with expected signs and have AMEs above 5% during banking crises. However, some variables (NCOTA, OREOTA, and TIETLB) during market crises and PD90TA during normal times become insignificant. For medium and large banks, the main result is that the ratio of total equity to total assets (TETA) remains significant with high AMEs at all times, primarily during banking crises. This is in line with Berger and Bouwman (2013) who find that higher capital improves the probability of surviving for medium and large banks during banking crises. Other findings among medium and large banks are relatively similar to the main results.

6 | CONCLUSION

The threat of bank failures influences not only the stability of the financial system but also the economy as a whole. For example, the failure of small banks in the early 1990s and the failure of large banks during the recent financial crisis are associated with considerable loan problems, profit reductions, credit risk, ineffective board of directors and their management, high unemployment, and low economic performance. Thus, a thorough analysis of such failures is central to policymakers, regulators, bank managers, and academics. Moreover, our results indicate meaningful institutional and policy implications. In effect, our findings emphasize the importance of considering bank size when designing appropriate policies and regulations targeted toward enhancing financial stability and resilience.

Although the literature has clarified the relevant drivers of bank failures, typically existing studies have not empirically analyzed the factors and the extent to which they are linked to the probability of bank failures across different size classes. In this study, we contribute to the extant literature by recognizing the differences in US bank failures engendered by size heterogeneity. We develop separate early-warning models for small, medium, and large banks, and report any differences in comparison to all bank failures prediction models, irrespective of bank size. We also compare the consistency (statistical significance and average marginal effects) of covariates when analysing bank failures across size categories. Furthermore, we contribute to the existing body of literature by using a statistical module building strategy suggested by Gupta et al. (2018) to develop parsimonious multivariate models from an exhaustive list of 61 accounting-based variables that have been employed significant predictors in existing bank failure literature.

The main empirical results show that factors associated with bank failures and the magnitudes of mutually significant factors (Average Marginal Effects) vary across small, medium, and large size categories. Further interesting results of this study are as follows. First, credit risk has a significant impact on bank failures probability across size classes and for the three lagged periods, implying that weak assets quality, represented by net charge off, past due 90+ days, loan loss reserves, and other real estate owned, increases the risk of failure. Second, small banks are most likely to fail if they have high deposit ratios, are more cost inefficient, and have a high liquidity risk, while medium and large banks with poor capital and low net interest margins are more likely to fail.

Our results are robust to up-to 3 years of lagged regression estimates, the inclusion of various control variables such as regulatory effects and house price inflation, interaction between bank size and bank charter, using an alternative bank size classification, and macroeconomic crisis periods and normal times. Moreover, the AUROC of all multivariate models developed across bank size classes for out-of-sample have an excellent performance for different forecast horizons.

This study has several interesting implications. As there is a lack of attention in the banking literature on the effects of factors on bank failures, the magnitude of these effects, and how they might differ across different size categories, it provides a thorough understanding of these issues. Given our findings that different factors have different effects on bank failures across bank size classes, researchers and policymakers developing early-warning models for predicting vulnerabilities leading to distress in banks should, whenever possible, take into consideration the differences in bank incentives engendered by size heterogeneity. Our approach, which develops separate early-warning models for small, medium, and large banks, and report any differences in comparison to all bank failures prediction models, irrespective of bank size, is one possibility to do so and can improve bank stability. Thus, our findings support bank regulators efforts to enhance the entire financial system.

The key limit to our study is the information content of market-based indicators. As the vast majority of commercial banks in the United States are not publicly traded, we focus on financial ratios based on accounting data. This limitation presents an opportunity for future work by using sample of publicly listed banks in developed and/or developing countries, and replicate our analysis supplemented with market-based measures.

DATA AVAILABILITY STATEMENT

I confirm that I have included a citation for available data in my references section.

ACKNOWLEDGMENTS

We thank the editor and anonymous referees for helpful comments and suggestions that have improved this paper significantly. We also acknowledge useful suggestions and comments of participants at the 17th Annual Conference of the European Economics and Finance Society (London, 2018), and 50th Money Macro and Finance Annual Conference (Edinburgh, 2018). An earlier version of this paper was circulated under the title “Bank size and bank failure”.

ENDNOTES

[1] Berger and Bouwman (2013) is an exception. They examine the impact of capital on bank performance (survival and market share) and how this effect differs among bank size classes (small, medium, and large) and across banking crises, market crises, and normal times. However, they focus only on one of the six CAMELS components that may misclassify distressed banks.

[2] However, we use their classification to report the robustness of our findings.

[3] The total of the count of banks across respective size categories is higher than the total number of banks in our sample due to the dynamic nature of banks' total assets. A bank may start small, but eventually move to the medium or large size categories as its total assets increases, or vice versa. For instance, a bank which is classified as small in 1990 may be classified as medium or large in 1995 due to increased asset size and vice-versa. Thus, some banks may appear in more than one size categories, but in different time periods.

[4] For more information about other authors, see Column 6 in Table A1.

[5] While calculating the financial ratios, zero values for all bank-year observations are replaced with $1 to avoid missing values.

[6] Summary statistics and correlation tables are not reported to save space, but are available from the authors upon request. A summary discussion on them is as follow: the mean of covariates bearing a positive relationship with bank failures (e.g., PD90TA) is expected to be higher for the failed group of banks than for its non-failed counterpart, and vice-versa. Contrarily, TETA, for instance, is expected to have a negative relationship with bank failures, and its mean across all size categories show that its value is lower for the failed group of observations than for its non-failed counterpart. Our expectations are well supported by all covariates across respective size categories except TDTA. The mean of TDTA for failed groups of banks is higher than for its non-failed counterpart, implying that failed banks have higher total deposits. This is

contrary to the intuition, where we expect failed banks to have funding and liquidity problems, and hence lower total deposits. Generally, median values of respective covariates reported are also sufficiently close to their respective mean values, thus problems that could arise due to significant skewness are not expected. Additionally, there is no unexpected variability in the values of standard deviation, minimum, and maximum for all variables across different bank size categories. The correlation matrix shows that some of the variables exhibit moderate to strong correlation with other variables. Issues associated with multicollinearity therefore has been addressed carefully when developing multivariate models.

[7] Except NIETA, which is insignificantly negative for 2-years lagged estimate and positive for 3-years lagged time without control variables.

[8] An exception is the coefficient of the 1-year lagged time without control variables, which is significantly positive.

[9] State chartered banks were subject to depositor preference law (DPL), which changes the priority structure of debt claims, from 1909, whereas nationally chartered banks were subject to DPL from 1993 onwards.

REFERENCES

Acharya, V., & Naqvi, H. (2012). The seeds of a crisis: A theory of bank liquidity and risk taking over the business cycle. *Journal of Financial Economics*, *106*(2), 349–366.

Aiyar, S., Calomiris, C. W., & Wieladek, T. (2014). Does macro-prudential regulation leak? Evidence from a UK policy experiment. *Journal of Money, Credit and Banking*, *46*(s1), 181–214.

Angbazo, L. (1997). Commercial bank net interest margins, default risk, interest-rate risk, and off-balance sheet banking. *Journal of Banking & Finance*, *21*(1), 55–87.

Arena, M. (2008). Bank failures and bank fundamentals: A comparative analysis of Latin America and East Asia during the nineties using bank-level data. *Journal of Banking & Finance*, *32*(2), 299–310.

Berger, A. N., & Bouwman, C. H. S. (2009). Bank liquidity creation. *Review of Financial Studies*, *22*(9), 3779–3837.

Berger, A. N., & Bouwman, C. H. S. (2013). How does capital affect bank performance during financial crises? *Journal of Financial Economics*, *109*(1), 146–176.

Berger, A. N., DeYoung, R., Genay, H., & Udell, G. F. (2000). Globalization of financial institutions: Evidence from cross-border banking performance. *Brookings-Wharton Papers on Financial Services*, *2000*(1), 23–120.

Berger, A. N., Imbierowicz, B., & Rauch, C. (2016). The roles of corporate governance in bank failures during the recent financial crisis. *Journal of Money, Credit and Banking*, *48*(4), 729–770.

Bertay, A. C., Demirgüç-Kunt, A., & Huizinga, H. (2013). Do we need big banks? Evidence on performance, strategy and market discipline. *Journal of Financial Intermediation*, *22*(4), 532–558.

Betz, F., Oprică, S., Peltonen, T. A., & Sarlin, P. (2014). Predicting distress in European banks. *Journal of Banking & Finance*, *45*, 225–241.

Bhagat, S., Bolton, B., & Lu, J. (2015). Size, leverage, and risk-taking of financial institutions. *Journal of Banking & Finance*, *59*, 520–537.

Bryant, J. (1980). A model of reserves, bank runs, and deposit insurance. *Journal of Banking & Finance*, *4*(4), 335–344.

Calomiris, C. (2007. November). *Bank failures in theory and history: The great depression and other 'contagious' events* (NBER Working Paper No. 13597). Cambridge, MA: National Bureau of Economic Research. Retrieved from https://www.nber.org/papers/w13597

Chernykh, L., & Cole, R. (2015). How should we measure bank capital adequacy for triggering prompt corrective action? A (simple) proposal. *Journal of Financial Stability*, *20*, 131–143.

Climent, F., Momparler, A., & Carmona, P. (2019). Anticipating bank distress in the Eurozone: An extreme gradient boosting approach. *Journal of Business Research*, *101*, 885–896.

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

Cole, R. A., & White, L. J. (2012). Déjà Vu all over again: The causes of U.S. commercial bank failures this time around. *Journal of Financial Services Research*, *42*(1–2), 5–29.

Cole, R. A., & Wu, Q. (2009). *Predicting bank failures using a simple dynamic hazard model*. Paper presented at 22nd Australasian Finance and Banking Conference. December. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.533.2611&rep=rep1&type=pdf

Danisewicz, P., McGowan, D., Onali, E., & Schaeck, K. (2017). Debt priority structure, market discipline, and Bank conduct. *The Review of Financial Studies*, *31*(11), 4493–4555.

De Haan, J., & Poghosyan, T. (2012). Size and earnings volatility of US bank holding companies. *Journal of Banking & Finance*, *36*(11), 3008–3016.

De Haas, R., Ferreira, D., & Taci, A. (2010). What determines the composition of banks' loan portfolios? Evidence from transition countries. *Journal of Banking & Finance*, *34*(2), 388–398.

De Jonghe, O., Diepstraten, M., & Schepens, G. (2015). Banks' size, scope and systemic risk: What role for conflicts of interest? *Journal of Banking & Finance*, *61*, S3–S13.

De Nicolo, G. (2000). *Size, charter value and risk in banking: An international perspective* (International Finance Discussion Paper 689). Board of Governors of the Federal Reserve System.

Demsetz, R. S., & Strahan, P. E. (1997). Diversification, size, and risk at bank holding companies. *Journal of Money, Credit and Banking*, *29*(3), 300–313.

Demyanyk, Y., & Hasan, I. (2010). Financial crises and bank failures: A review of prediction methods. *Omega*, *38*(5), 315–324.

DeYoung, R. (1998). Management quality and X-inefficiency in National Banks. *Journal of Financial Services Research*, *13*(1), 5–22.

DeYoung, R. (2003). The failure of new entrants in commercial banking markets: A split-population duration analysis. *Review of Financial Economics*, *12*(1), 7–33.

DeYoung, R., & Torna, G. (2013). Nontraditional banking activities and bank failures during the financial crisis. *Journal of Financial Intermediation*, *22*(3), 397–421.

Diamond, D. W., & Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*, *91*(3), 401–419.

Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its Oracle properties. *Journal of the American Statistical Association*, *96*(456), 1348–1360.

Gupta, J., Gregoriou, A., & Ebrahimi, T. (2018). Empirical comparison of hazard models in predicting SMEs failure. *Quantitative Finance*, *18*(3), 437–466.

Hakenes, H., & Schnabel, I. (2011). Bank size and risk-taking under Basel II. *Journal of Banking & Finance*, *35*(6), 1436–1449.

Haslem, J. A., Scheraga, C. A., & Bedingfield, J. P. (1992). An analysis of the foreign and domestic balance sheet strategies of the U.S. banks and their association to profitability performance. *Management International Review*, *32*, 55–75.

Hong, H., Huang, J.-Z., & Wu, D. (2014). The information content of Basel III liquidity risk measures. *Journal of Financial Stability*, *15*, 91–111.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). New Jersey: John Wiley & Sons.

Imbierowicz, B., & Rauch, C. (2014). The relationship between liquidity risk and credit risk in banks. *Journal of Banking & Finance*, *40*, 242–256.

Jin, J. Y., Kanagaretnam, K., & Lobo, G. J. (2011). Ability of accounting and audit quality variables to predict bank failure during the financial crisis. *Journal of Banking & Finance*, *35*(11), 2811–2819.

Keeley, M. C. (1990). Deposit insurance, risk, and market power in banking. *The American Economic Review*, *80*, 1183–1200.

Khan, M. S., Scheule, H., & Wu, E. (2017). Funding liquidity and bank risk taking. *Journal of Banking & Finance*, *82*, 203–216.

Köhler, M. (2015). Which banks are more risky? The impact of business models on bank stability. *Journal of Financial Stability*, *16*, 195–212.

Kolari, J., Glennon, D., Shin, H., & Caputo, M. (2002). Predicting large US commercial bank failures. *Journal of Economics and Business*, *54*(4), 361–387.

Laeven, L., Ratnovski, L., & Tong, H. (2016). Bank size, capital, and systemic risk: Some international evidence. *Journal of Banking & Finance*, *69*, S25–S34.

Lane, W. R., Looney, S. W., & Wansley, J. W. (1986). An application of the cox proportional hazards model to bank failure. *Journal of Banking & Finance*, *10*(4), 511–531.

Liu, W.-M., & Ngo, P. T. H. (2014). Elections, political competition and bank failure. *Journal of Financial Economics*, *112*(2), 251–268.

Loutskina, E. (2011). The role of securitization in bank liquidity and funding management. *Journal of Financial Economics*, *100*(3), 663–684.

Meyer, P. A., & Pifer, H. W. (1970). Prediction of bank failures. *The Journal of Finance*, *25*(4), 853–868.

Ng, J., & Roychowdhury, S. (2014). Do loan loss reserves behave like capital? Evidence from recent bank failures. *Review of Accounting Studies*, *19*(3), 1234–1279.

Pais, A., & Stork, P. A. (2013). Bank size and systemic risk. *European Financial Management*, *19*(3), 429–451.

Poghosyan, T., & Čihak, M. (2011). Determinants of bank distress in Europe: Evidence from a new data set. *Journal of Financial Services Research*, *40*(3), 163–184.

Repullo, R. (2004). Capital requirements, market power, and risk-taking in banking. *Journal of Financial Intermediation*, *13*(2), 156–182.

Schaeck, K. (2008). Bank liability structure, FDIC loss, and time to failure: A quantile regression approach. *Journal of Financial Services Research*, *33*(3), 163–179.

Thomson, J. B. (1992). Modeling the bank regulator's closure option: A two-step logit regression approach. *Journal of Financial Services Research*, *6*(1), 5–23.

Tian, S., Yu, Y., & Guo, H. (2015). Variable selection and corporate bankruptcy forecasts. *Journal of Banking and Finance*, *52*, 89–100.

Wheelock, D. C., & Wilson, P. W. (2000). Why do banks disappear? The determinants of U.S. bank failures and acquisitions. *Review of Economics and Statistics*, *82*(1), 127–138.

How to cite this article: Alzugaiby B, Gupta J, Mullineux A, Ahmed R. Relevance of size in predicting bank failures. *Int J Fin Econ*. 2021;26: 3504–3543. https://doi.org/10.1002/ijfe.1973

APPENDIX A.

TABLE A1 Description of variables

No. (1)	Category (2) Bank size	Variable (3) LTA	Description (4) Natural logarithm of total assets	CALL item codes (5) rcfd2170	Source (6) (Bertay et al., 2013)
Explanatory variables					
1	Capital (C)	TETA	Total equity divided by total assets	rcfd3210/ rcfd2170	(Berger et al., 2016)
2		T1CR	Tier1 capital ratio	rcfd7206	(Betz et al., 2014)
3		NPACR	Nonperforming assets coverage ratio = [(Equity + LLR) - Weighted NPA] divided by total assets	[(rcfd3210+ rcfd3123) - (0.20*rcfd1406 + 0.50*rcfd1407+ 1*(rcfd1403 + rcfd2150)]/ rcfd2170	(Chernykh & Cole, 2015)
4	Asset quality (A)	LLRTA	Loan loss reserves divided by total assets	rcfd3123/rcfd2170	(Cole & White, 2012)
5		PD90TA	Loans past due 90+ days divided by total assets	rcfd1407/rcfd2170	(Cole & Wu, 2009)
6		NAATA	Nonaccrual loans divided by total assets	rcfd1403/rcfd2170	(Cole & Wu, 2009)
7		OREOTA	Other real estate owned divided by total assets	rcfd2150/rcfd2170	(Cole & Wu, 2009)
8		NPATA	Non-performing assets (PD38-89 + PD90 + Nonaccrual loans + Other real estate owned) divided by total assets	(rcfd1406 + rcfd1407+ rcfd1403 + rcfd2150)/ rcfd2170	(Cole & White, 2012)
9		LLRNPL	Loan loss reserves divided by non-performing loans	rcfd3123/rcfd2170	
10		LLPTL	Loan loss provisions divided by total loans	riad4230/rcfd2122	(Poghosyan & Čihak, 2011)
11		LLPTA	Loan loss provisions divided by total assets	riad4230/rcfd2170	(Kolari et al., 2002)
12		NPLTL	Non-performing loans divided by total loans	(rcfd1407 + rcfd1403)/rcfd2122	(Danisewicz et al., 2017)
13		NPLTA	Non-performing loans divided by total assets	(rcfd1407 + rcfd1403)/rcfd2170	(Wheelock & Wilson, 2000)
14		NCOTA	Net-charge offs divided by total assets	(riad4635-riad4605)/rcfd2170	(Kolari et al., 2002)
15		RELTA	Real estate loans divided by total assets	rcfd1410/rcfd2170	(Ng &
		Roychowdhury, 2014)	16		CILTA
	Commercial & industrial loans divided by total asset	rcfd1766/rcfd2170	(Cole & White, 2012)		
17		CLTA	Consumer loans divided by total asset	rcfd1975/rcfd2170	(Cole & White, 2012)
18		CDLTA	Construction & development loans divided by total assets	rcon1415/rcfd2170	(Cole & White, 2012)
19		RERLTA	Real estate residential (1–4) family loans divided by total assets	rcon1430/rcfd2170	(Cole & White, 2012)

(Continues)

TABLE A1 (Continued)

No. (1)	Category (2)	Variable (3)	Description (4)	CALL item codes (5)	Source (6)
	Bank size	**LTA**	**Natural logarithm of total assets**	**rcfd2170**	**(Bertay et al., 2013)**
20		REMLTA	Real estate residential multifamily loans divided by total assets	rcon1460/rcfd2170	(Cole & White, 2012)
21		RENFNRLTA	Real estate nonfarm non-residential loans divided by total assets	rcon1480/rcfd2170	(Cole & White, 2012)
22	Management (M)	NIETA	Cost inefficiency; noninterest expenses divided by total assets	riad4093/rcfd2170	(DeYoung & Torna, 2013)
23		CIR	Cost to income ratio; operating expenses divided by operating income	riad4130/riad4000	(Poghosyan & Čihak, 2011)
24	Earnings (E)	NIM	Net interest margin; net interest income divided by average earning assets	riad4074/rcfd3402	(Betz et al., 2014)
25		ROA	Return on assets; net income divided by total assets	riad4340/rcfd2170	(Arena, 2008)
26		ROE	Return on equity; net income divided by total equity	riad4340/rcfd3210	(Berger & Bouwman, 2013)
27	Liquidity (L)	CDTA	Cash & due divided by total asset	rcfd0010/rcfd2170	(Cole & White, 2012)
28		TSTA	Total securities divided by total assets	rcfd8641/rcfd2170	(Cole & White, 2012)
29		TLTA	Total loans divided by total assets	rcfd2122/rcfd2170	(Kolari et al., 2002)
30		LATLB	Liquid assets divided by total liabilities	[rcfd0010 + (rcfd0390 & rcfd1773 + rcfd1754)]/ rcfd2948	(Poghosyan & Čihak, 2011)
31		LATA	Liquid assets (Cash & due from banks + securities held for investment + securities held for sale) divided by total assets	[rcfd0010 + (rcfd0390 & rcfd1773 + rcfd1754)]// rcfd2170	(Arena, 2008)
32		FTA	(Fed fund purchase - fed fund sold) divided by total assets	(rcfd2800-rcfd1350)/rcfd2170	(Wheelock & Wilson, 2000)
33		TRADTA	Trading asset divided by total assets	rcfd3545/rcfd2170	(DeYoung & Torna, 2013)
34		TIETLB	Total interest expenses divided by total liabilities	riad4073/rcfd2948	(Danisewicz et al., 2017)
35	Sensitivity to market	TIOI	Trading income divided by operating income	riada220/riad4000	(Betz et al., 2014)
36	Funding	TDTA	Total deposits divided by total assets	rcfd2200/rcfd2170	(Acharya & Naqvi, 2012)
37		STDTD	Short-term deposits (transaction deposits + demand deposits) divided by total deposits	(rcon2215 + rcon2210)/rcfd2200	(Berger et al., 2016)
38		BDTA	Brokered deposits divided by total assets	rcon2365/rcfd2170	(Cole & White, 2012)
39		LCDTA	Large certificates of deposits ($100,000 & more) divided by total assets	rcon2604/rcfd2170	(Cole & Wu, 2009)

TABLE A1 (Continued)

No. (1)	Category (2)	Variable (3)	Description (4)	CALL item codes (5)	Source (6)
	Bank size	**LTA**	**Natural logarithm of total assets**	**rcfd2170**	**(Bertay et al., 2013)**
40		LCDTLB	Large certificates of deposits divided by total liabilities	rcon2604/rcfd2948	(Cole & Wu, 2009)
41		MBSTA	Mortgage-backed securities divided by total assets	rcfd8639/rcfd2170	(Cole & White, 2012)
42	Business model	NDFTLB	Non-deposit funding divided by total liabilities	rcfd2527/rcfd2948	(Köhler, 2015)
43		NIIOI	Non-interest income divided by operating income	riad4079/riad4000	(Bertay et al., 2013)
44	Leverage	TATE	Total assets divided by total equity	rcfd2170/rcfd3210	
45		TLBTE	Total liabilities divided by total equity	rcfd2950/rcfd3210	(Betz et al., 2014)
46		TLBTA	Total liabilities divided by total assets	rcfd2948/rcfd2170	(Danisewicz et al., 2017)
47		TLTD	Total loans divided by total deposits	rcfd2122/rcfd2200	(Betz et al., 2014)
48	Growth	GTA	Growth of total assets		(Cole & White, 2012)
49		GTL	Growth of total loans		(Berger et al., 2016)
50	Other	GWTA	Goodwill divided by total assets	rcfd3163/rcfd2170	(Cole & White, 2012)
51		LIR	Loans interest rate; total interest income divided by total loans	riad4107/rcfd2122	(Arena, 2008)
52	Market discipline	DIR	Deposits interest rate; total interest expense divided by total deposits	riad4073/rcfd2200	(Arena, 2008)
53		SPREAD	LIR – DIR		(Arena, 2008)
54	Non-traditional	ICFTA	Insurance commissions and fees divided by total assets	riadb494/rcfd2170	(DeYoung & Torna, 2013)
55		IRUITA	Insurance & reinsurance underwriting income divided by total assets	riadc386/rcfd2170	(DeYoung & Torna, 2013)
56		VCRTA	Venture capital revenue divided by total assets	riadb491/rcfd2170	(DeYoung & Torna, 2013)
57	FCSBTA	Fees & commissions from securities brokerage divided by total assets	riadc886/rcfd2170	(DeYoung & Torna, 2013)	
58		NSITA	Net securitization income divided by total assets	riadb493/rcfd2170	(DeYoung & Torna, 2013)
59		IBFCTA	Investment banking fees & commissions divided by total assets	riadb490/rcfd2170	(DeYoung & Torna, 2013)
60		NSFTA	Net servicing fees divided by total assets	riadb492/rcfd2170	(DeYoung & Torna, 2013)
61	Off balance sheet	TUCTA	Total unused commitment divided by total assets.	rcfd3423/rcfd2170	(Berger et al., 2016)

(Continues)

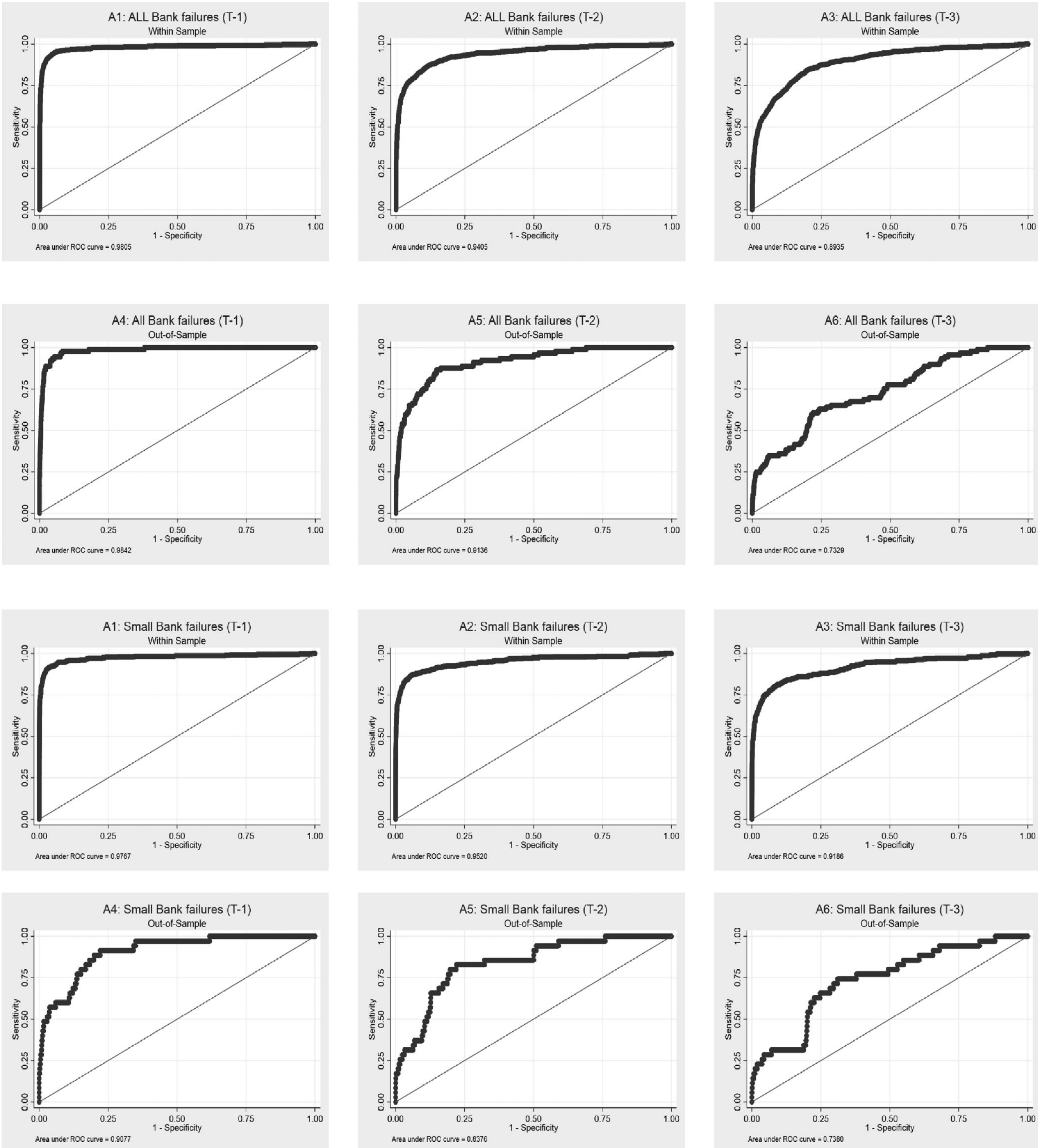
TABLE A1 (Continued)

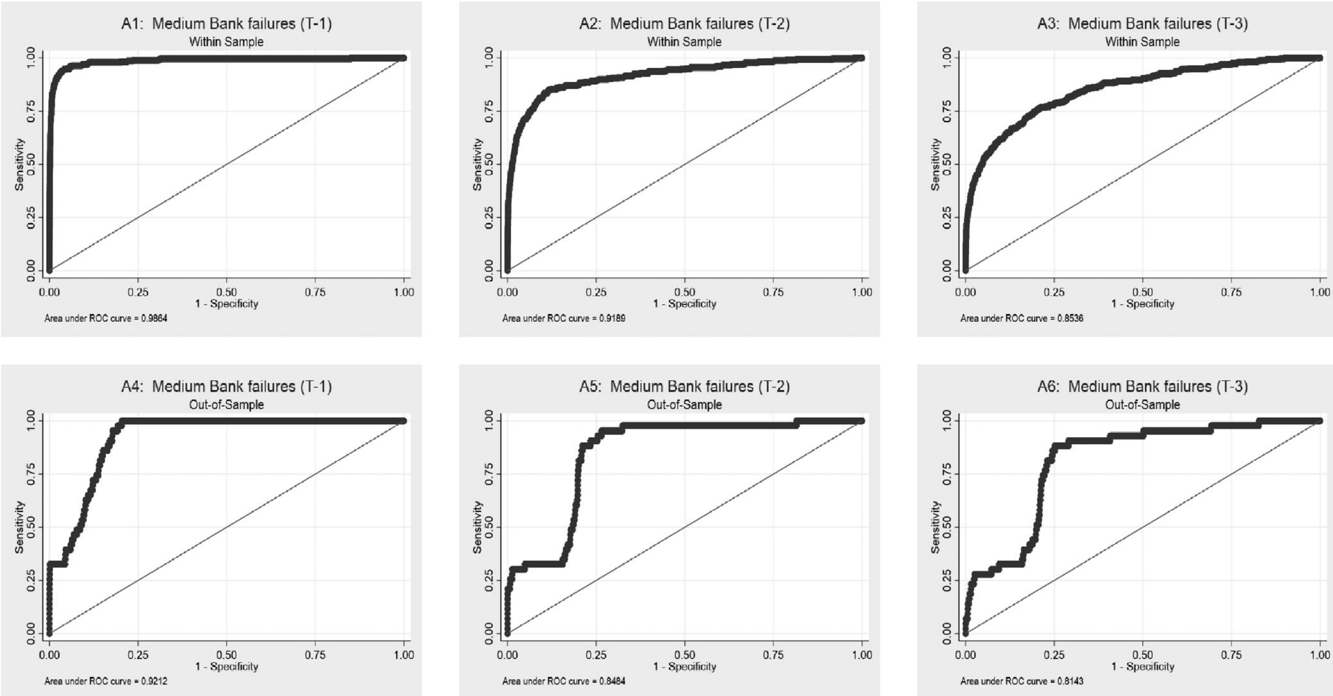
No. (1)	Category (2)	Variable (3)	Description (4)	CALL item codes (5)	Source (6)
	Bank size	**LTA**	**Natural logarithm of total assets**	**rcfd2170**	**(Bertay et al., 2013)**
Control variables					
62	Primary regulators	FDIC	Dummy variable indicating whether the bank is a state-chartered and non-member of the Federal Reserve System.		(Berger & Bouwman, 2013)
63		FED	Dummy variable indicating whether the bank is a state-chartered and member of the Federal Reserve System.		(Berger & Bouwman, 2013)
64	Foreign ownership	FOPCT	Dummy variable indicating whether the bank is foreign-owned (25% or more).		(Berger & Bouwman, 2013)
65	Growth of house prices index	GHPI	State-level house price indices (HPIs) of the seasonally adjusted Federal Housing Finance Agency's (FHFA).		(Berger & Bouwman, 2013)
66	Banking crises	SL	Dummy variable indicating whether the year is on saving and loans crisis that occurred between 1987 and 1990.		(Berger & Bouwman, 2013)
67		GFC	Dummy variable indicating whether the year is on subprime lending crisis (Global financial crisis) that occurred between 2008 and 2010.		(Berger & Bouwman, 2013)

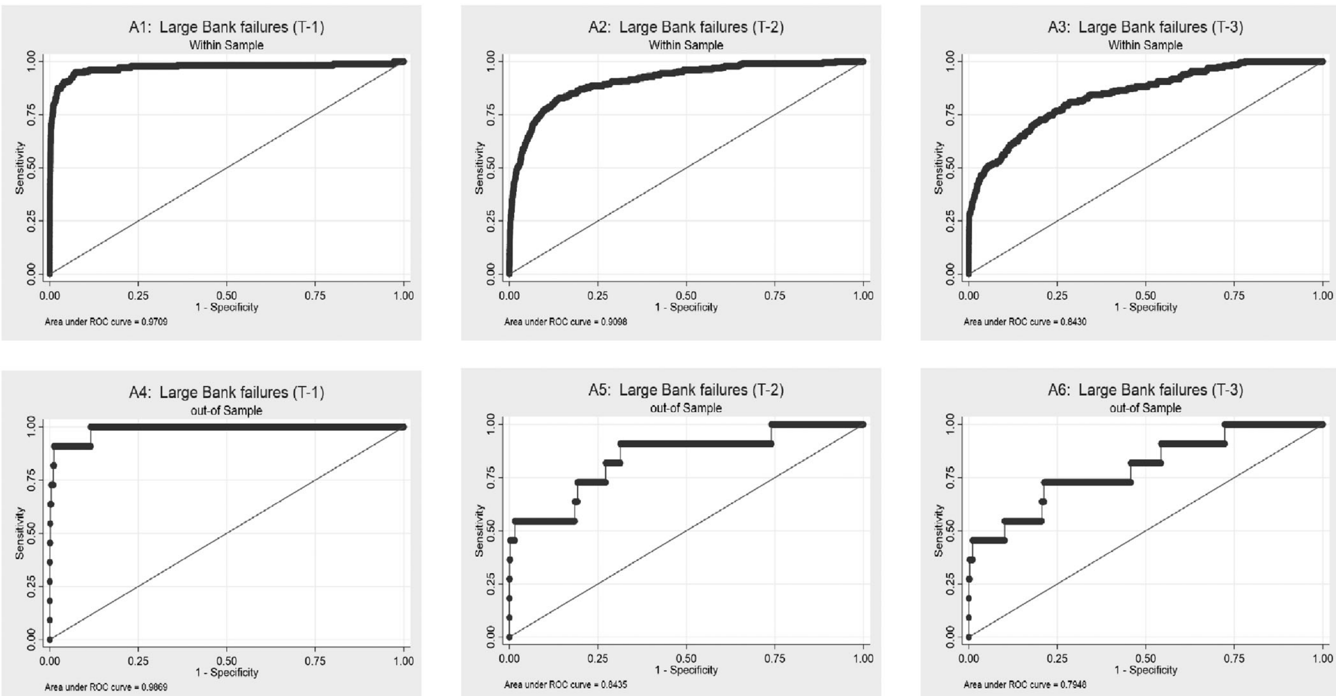
Notes: This table reports the set of explanatory and control variables that we use in our empirical analysis. The first column is the number of explanatory and control variables, while the second column lists the category of explanatory and control variables. The third column lists names of variables. The fourth column provides their respective definitions. Financial information is obtained from the Call Report (FDIC) database, covering an analysis period from 1985 to 2016. The last column states the specific codes of Call Report data items that we use to calculate explanatory variables.

A.1. | Table of area under ROC curves

The receiver operating characteristics (ROC) curve and the area under the ROC (AUROC) curve are non-parametric measures to evaluate the classification performance of early-warning models developed to identify distressed banks (Betz et al., 2014). The ROC curve describes the trade-off between true-positive (sensitivity: a bank actually fails, and the model classifies it as expected failure) and false-negative (1 – specificity: a bank actually fails but the model classifies it as expected survival) for an entire range of classification thresholds (Gupta et al., 2018). However, ROC offers a range of performance assessments. This means that the accuracy of the predicted class probabilities is mostly overlooked. We therefore use the AUROC, which is by far the most common non-parametric method for evaluating a fitted prediction model's ability to assign a randomly chosen positive instance higher than a randomly chosen negative one (Betz et al., 2014; Cole & White, 2012). In other words, the AUROC gauges the ability of the prediction model to discriminate between those banks which experience the event of interest, and those which do not. Its value varies between 0.5 and 1.0, which summarizes the classification performance of the model developed. The value of 1 represents a perfect model, whereas the value of 0.5 represents no discrimination ability of the model. Generally, there is no guide for classifying the predictive accuracy of a model based on AUROC, however any value above 0.7 is acceptable and above 0.8 is considered to be excellent (Hosmer et al., 2013). Thus, the higher the AUROC, the better the model's prediction performance. Although few studies (e.g., Betz et al., 2014; Poghosyan & Čihak, 2011) in the literature of bank failures have reported the AUROC, from a policy perspective and for the empirical tests in this paper this metric is fundamental for comparing performance and providing a validation of the models. Following the approach of Gupta et al., (2018), we report area under ROC (AUROC) curves for respective models to evaluate the within-sample and out-of-sample classification performance of the models developed. For within-sample validation, we estimate the models using the entire sample data. To validate models' out-of-sample predictive performance, we first estimate the models using all available information up to the year 2011, and then predict the probability of bank failures for the year 2012. Subsequently, we incorporate 2012 in the estimation sample and predict the probability of bank failures for 2013 and so on, up to the year 2016. Finally, we use these predicted probabilities from the year 2012 until the year 2016 to estimate out-of-sample AUROC for respective multivariate regression models.


A1: ALL Bank failures (T-1)
Within Sample
Sensitivity
1 - Specificity
Area under ROC curve = 0.9805
A2: ALL Bank failures (T-2)
Within Sample
Area under ROC curve = 0.9405
A3: ALL Bank failures (T-3)
Within Sample
Area under ROC curve = 0.8935
A4: All Bank failures (T-1)
Out-of-Sample
Area under ROC curve = 0.9842
A5: All Bank failures (T-2)
Out-of-Sample
Area under ROC curve = 0.9136
A6: All Bank failures (T-3)
Out-of-Sample
Area under ROC curve = 0.7329
A1: Small Bank failures (T-1)
Within Sample
Area under ROC curve = 0.9767
A2: Small Bank failures (T-2)
Within Sample
Area under ROC curve = 0.9520
A3: Small Bank failures (T-3)
Within Sample
Area under ROC curve = 0.9186
A4: Small Bank failures (T-1)
Out-of-Sample
Area under ROC curve = 0.9077
A5: Small Bank failures (T-2)
Out-of-Sample
Area under ROC curve = 0.8376
A6: Small Bank failures (T-3)
Out-of-Sample
Area under ROC curve = 0.7380


A1: Medium Bank failures (T-1)
Within Sample
Area under ROC curve = 0.9864
A2: Medium Bank failures (T-2)
Within Sample
Area under ROC curve = 0.9189
A3: Medium Bank failures (T-3)
Within Sample
Area under ROC curve = 0.8536
A4: Medium Bank failures (T-1)
Out-of-Sample
Area under ROC curve = 0.9212
A5: Medium Bank failures (T-2)
Out-of-Sample
Area under ROC curve = 0.8484
A6: Medium Bank failures (T-3)
Out-of-Sample
Area under ROC curve = 0.8143
Sensitivity
1 - Specificity
0.00
0.25
0.50
0.75
1.00


A1: Large Bank failures (T-1)
Within Sample
Area under ROC curve = 0.9709
A2: Large Bank failures (T-2)
Within Sample
Area under ROC curve = 0.9098
A3: Large Bank failures (T-3)
Within Sample
Area under ROC curve = 0.8430
A4: Large Bank failures (T-1)
out-of Sample
Area under ROC curve = 0.9869
A5: Large Bank failures (T-2)
out-of Sample
Area under ROC curve = 0.8435
A6: Large Bank failures (T-3)
out-of Sample
Area under ROC curve = 0.7948
Sensitivity
1 - Specificity
0.00
0.25
0.50
0.75
1.00