Journal of Banking & Finance 64 (2016) 101–111

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

An efficient and functional model for predicting bank distress: In and out of sample evidence ☆

Sean Cleary [a,*], Greg Hebb [b,1]

[a] Queen's School of Business, Queen's University, Kingston, Ontario K7L 3N6, Canada
[b] Rowe School of Business, Dalhousie University, Halifax, Nova Scotia B3H 4R2, Canada

ARTICLE INFO

Article history:
Received 23 December 2014
Accepted 12 December 2015
Available online 21 December 2015

JEL classification:
G21
G33

Keywords:
Bank failures
Financial distress
Financial crisis

ABSTRACT

We examine the failures of 132 U.S. banks over the 2002–2009 period using discriminant analysis and successfully distinguish between banks that failed and those that didn't 92% of the time using in-sample quarterly data. Our two most important variables are related to bank capital and loan quality, as one might expect; although bank profitability is also important. The resulting model is then used out-of-sample to examine the failure of 191 banks during 2010–11, with predictive accuracy in the 90–95% range.

Our results demonstrate that our model can also easily be applied to a large number of firms (even those that don't fail) and does an excellent job of distinguishing healthy from distressed banks. Combining this effectiveness with its ease of implementation makes it very functional. Such a model should be of obvious interest to regulators, analysts, and all those with a direct interest in assessing bank financial health.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

There have been numerous studies attempting to predict business failures, dating back to the 1930s. Altman (1968) is one of the most widely cited studies based on his success in predicting the bankruptcy of manufacturing firms using discriminant analysis. Numerous other methodologies have been used throughout the years including logit models, probit models, neural networks, and hazard rate models. We make use of the discriminant analysis approach, which has been shown to be reliable and easy to use through the years, to examine U.S. bank failures over the 2002–2011 period.

Our study is very timely given the recent financial crisis, and we are able to make a significant contribution for several reasons. First, there have been a large number of bank failures and banks in distress that have required regulatory intervention in the global financial sector since 2008. Focusing on U.S. bank failures has provided us with an interesting and sizable sample of firms to work with. While the number of distressed banks in the U.S. has declined steadily since 2010, this has not been the case in Europe where banks in many countries are still facing extreme pressures. All of this suggests the critical importance of this topic. While there have been several recent related studies, ours stands out for its effectiveness, and for its ease of implementation. Second, given the widespread fallout from bank failures to the financial system and to the U.S. and global economies, much attention has been focused by economists, politicians, regulators, etc. as to how to put measures in place to avoid such occurrences happening again in the future. At the very least, all interested parties want to be able to better recognize the warning signs further in advance – far enough in advance to take appropriate actions. Our model can be easily applied to a large number of banks and thus provides very useful information for interested parties regarding banks that are in serious risk of failing, those that should be reviewed for risk rating "downgrades," and those that are prime candidates for "on-site" examinations, such as those conducted by the Federal Deposit Insurance Corporation (FDIC). It can also be employed to assess the overall level of health of the financial system, and to monitor any significant changes in the system. Thirdly, our results

☆ The authors would like to acknowledge financial support from the Social Sciences Humanities and Research (SSHRC) Canada, and the Smith School of Business at Queen's University. We are grateful for comments received from participants at the 2014 World Finance Conference, and at a finance seminar at the Schulich School of Business at York University. The paper was improved substantially by incorporating the suggestions of an anonymous referee and Associate Editor, and from the Managing Editor, Carol Alexander.

* Corresponding author. Tel.: +1 613 533 3384.
E-mail addresses: scleary@business.queensu.ca (S. Cleary), Gregory.hebb@dal.ca (G. Hebb).
[1] Tel.: +1 902 494 1802.

http://dx.doi.org/10.1016/j.jbankfin.2015.12.001
0378-4266/© 2015 Elsevier B.V. All rights reserved.

demonstrate the critical importance of using the most "timely" available data at any point in time, which reflects the fact that conditions can deteriorate very quickly for banks that find themselves facing distress. Finally, our model may be easily applied to other markets, and we plan to do so in our future work.

We use a recent and comprehensive sample obtained from the Bankscope database to examine the failures of 132 U.S. banks over the 2002–2009 period, after constructing a matched sample of non-failed banks for comparison purposes. The resulting model that is determined using these in-sample observations is then used out-of-sample to examine the failure of 191 banks during 2010–11. We use multivariate discriminant analysis (MDA) to predict banks that would fail using three groups of variables: (1) Financial Health Variables – variations of those used by the original Altman (1968) study related to financial health (suitably modified to apply to banks rather than manufacturing firms); (2) On Balance Sheet Control Variables – we use several variables to control for liquidity, reliance on loans, loan quality and capital adequacy; and, (3) Off Balance Sheet Control Variables – to measure for the amount of off balance sheet items, acceptances, guarantees, etc.

Our model successfully distinguishes between banks that would fail and those that wouldn't 92% of the time using our in-sample quarterly data. This type of in-sample prediction success exceeds that of most related studies. We also found that the use of quarterly data greatly improved our results, when compared to the use of less timely annual data[2]. This is an important observation that is consistent with Cole and Gunther (1998)'s observation that information used to assess bank financial health can become obsolete within six months. This highlights the importance of regular monitoring, particularly for banks that appear to be troubled. Our two most important variables are related to bank capital and loan quality, as one might expect; although bank profitability is also important. Even more important than our in-sample prediction success is the fact that our model does an excellent job of distinguishing healthy banks from those that are at high risk of failure (even those that do not fail), in terms of traditional variables that have been shown to be important to bank health. Our evidence suggests our model is extremely effective in doing so, and is easy to implement on a large sample of banks using available data.

Finally, when we apply our model "out-of-sample" in 2010–11, the predictive accuracy is in the 90–95% range, and the model also provides very effective assessments of bank health for a large sample of banks. This out-of-sample success is superior to most related studies and is a clear indication of the effectiveness of the model. Combining this effectiveness with its ease of implementation makes this a very attractive and functional model that will be of obvious interest to banks' internal risk management, to regulators, as well as to others with a direct interest in assessing bank financial health such as analysts, fund managers, and financial transaction counter-parties.

While a statistical model such as ours is no substitute for "on-site" examinations and detailed qualitative analysis, it does represent an effective supplement to them. In particular, since we are able to effectively classify a large number of firms, it can be used to identify those banks that should be scrutinized in more detail, and hence make use of more detailed qualitative information.

The remainder of the paper is organized as follows: Section 2 provides motivation for our study, Section 3 describes the data set and methodologies employed, Section 4 discusses our in-sample results, Section 5 discusses our out-of-sample predictions, while Section 6 concludes.

2. Motivation

2.1. Bank failures

Berger and Bouwman (2009) suggest that an abundant theoretical literature suggests that "banks exist because they perform two central roles in the economy – they create liquidity and they transform risk." In particular, we know that banks transform risk and create liquidity by financing themselves with highly liquid, low-risk deposits and investing in higher-risk, illiquid assets (mainly in the form of loans), as discussed in Boot and Thakor (2010), Bhattacharya and Thakor (1993), Diamond (1984), and Diamond and Dybvig (1983) for example. As a result of performing these basic functions, the average bank faces a significant amount of both insolvency and liquidity risk. Therefore, it is not surprising that we see a rise in the number of bank failures during periods where the market value of bank assets (i.e., loans) have deteriorated and/or when there is significant uncertainty regarding the stability of deposit bases.

The most recent U.S. banking crisis was precipitated by a significant deterioration in real estate prices, which caused the value of real estate loans and related debt instruments to decline sharply, leading to the failure of a large number of banks. Indeed, U.S. regulators shut down 30 banks in 2008 versus only 27 during the entire 2000–2007 period. The number of closures then increased to 140 in 2009 and to 162 in 2010, before declining to 92 in 2011, 51 in 2012, and 25 in 2013. These figures can be seen in Fig. 1.

Not surprisingly, the FDIC fund fell into the red during 2009. This represents its first negative balance since the savings and loans crisis in the early 1990s – in fact the fund had recorded surpluses of over $40 billion since 2000, exceeding $50 billion in 2006. During September of 2009, the FDIC proposed the unprecedented step of having the banking industry prepay $45 billion in fees by the end of the year to give the government more room to handle future failures. At the time they announced this decision, they had already used $30 billion to cover bank failures over the next year. Estimates at that time suggested that bank failures from 2009 to 2013 would cost the FDIC $100 billion. According to Bloomberg[3], actual failures slowed more than predicted and the fund actually had a positive balance of $11.8 billion by the end of 2011. This was still significantly below the required 1.15% of insured deposits and it is not predicted to reach this level until 2018.

In response to such pressures, during 2009, the FDIC increased the fees charged for deposit insurance[4]. For example, as of April 1, 2009 the "base assessment rates" (in basis points) ranged from 12 to 16 for Risk Category I banks, 22 for Category II banks, 32 for Category III banks, and 45 for Category IV banks, versus 2006 rates of 2–4, 7, 27 and 40, respectively. These rates were raised again in 2011 with a total range of 12 to 77 basis points depending on the risk category. These risk categories are determined by a bank's CAMELS (Capital, Asset quality, Management, Earnings, Liquidity and market Sensitivity) rating. This rating is based upon on-site bank examinations conducted by the FDIC. In addition to qualitative information, the rating is based upon various ratios including: the Tier 1 Capital ratio; (loans past due 30–89 days)/(gross assets); (nonperforming assets)/(gross assets); (net loan charge-offs)/(gross assets); and, (net income before taxes)/(risk-weighted assets).

In order to supplement on-site examinations, and to assist in prioritizing which banks are in most urgent need of such examinations, the FDIC also estimates a bank's stability using a Statistical

[2] The difference was much more pronounced in the out-of-sample analysis, which will be discussed later.

[3] *Source:* http://www.bloomberg.com/news/2012-04-23/fdic-says-deposit-insurance-fund-recovery-expected-by-late-2018.html

[4] *Source:* www.FDIC.gov, "Federal Regulator," Vol. 74 No. 41, March 4, 2009, Rules and Regulations.

U.S. Failed Banks (by year)

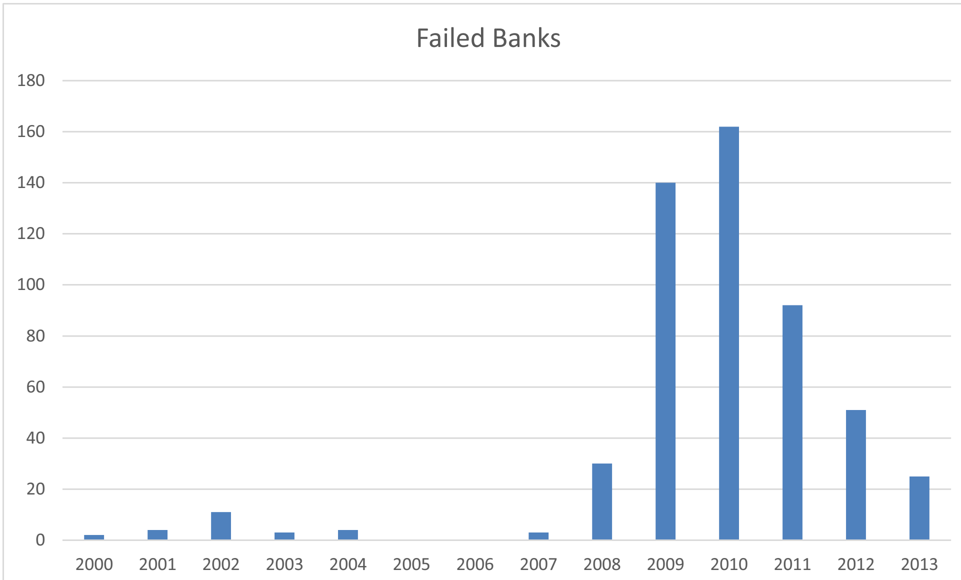

Fig. 1. U.S. failed banks (by year).

CAMELS Off-site Rating system (SCOR), which is described by Collier et al. (2005). This system uses twelve ratios related to the six CAMELS categories, and is re-estimated every quarter. Collier et al. (2005) suggest that the variables for the original specification were chosen based upon "both a review of the literature on bank failures and discussions with bank examiners."

The SCOR system is used to both estimate a bank's CAMELS rating, and to estimate the probability of a rating downgrade. Collier et al. (2005) examine the performance of this model over the 1986–2002 period and find that it accurately predicts only 27 percent of the banks that are subsequently downgraded in the following four-to-six month period. They suggest that despite this low accuracy, it is still "approximately nine times better than a random guess." In addition, they argue the model provides useful information by flagging areas of concern for the banks. They go on to point out that this system which is based purely on financial ratios is limited by design, but that it represents a useful "complement" to on-site examination, rather than a substitute for them. On-site visits in turn can be used to identify "unsafe practices before they affect the bank's financial condition" and "can also detect misstated financial reports." In essence, Collier et al. (2005) are confirming the usefulness of statistical models based on financial ratios, despite their inherent limitations.

2.2. Related empirical evidence[5]

Bellovary et al. (2007) provide a review of bankruptcy prediction studies from 1930 to 2004 related to general businesses, as well as financial institutions. They find various levels of success across the models in terms of predicting which firms will go bankrupt "in-sample" (from 20% to 100%). The models employed have tended to be primarily multivariate discriminant analysis (37%), logit analysis (21%), probit analysis (4%) and neural networks (23%). They go on to analyze the 165 studies since 1985 and note that "multivariate discriminant analysis and neural networks are the most promising methods." They also find that the average number of factors employed has averaged around 10, and that "higher model accuracy is not guaranteed with a greater number of factors." They note that "some models with two factors are just as capable of accurate prediction as models with 21 factors."

Bellovary et al. (2007)'s review includes 15 studies related to the bankruptcy of banks. Of these, three specifically used discriminant analysis to predict bank failure. Sinkey (1975) correctly predicted 72% of problem banks during the 1972–1973 time period. Pettway and Sinkey (1980) examined 33 banks that failed between 1970 and 1975. Using accounting data and discriminant analyses, they correctly classified between 75% and 82% of failed banks up to four years in advance. Finally, Espahbodi (1991) correctly predicted 86% of bank failures one year in advance and 84% two years in advance.

Subsequent to the sample period of Bellovary et al. (2007), there have been several studies of bank failure. In a study that pre-dates the recent financial crisis, Kosmidou and Zopounidis (2008) study U.S. bank failures over the 1993–2003 period and achieve 89% success with their model. Not surprisingly, given the large number of recent failures, and the attention drawn to the issue, several recent U.S. studies discussed below have focused on bank distress and failures during the financial crisis time period. We would note at the start of this discussion that a large majority of these studies use variables to proxy for the CAMELS characteristics discussed above.

Aubuchon and Wheelock (2010) study the relationship between regional economic factors and bank failure and find a significant correlation between these failures and state economic conditions. In particular, they find bank failure rates were higher in states that had larger housing price declines. Jin et al. (2011) examines the ability of accounting and auditor data to predict bank failure during the financial crisis. They use a probit model based on 2006 data to determine which factors are significant predictors of bank failure. While not including any out of sample predictions, they find auditor type, auditor industry specialization, Tier 1 capital ratio, proportion of securitized loans, growth in loans, and loan mix are

[5] In this section, we provide only a limited discussion of the variables used in each study for the sake of brevity, but would note that several of the same variables are used in many of them. Instead, in Section 3, we will reference the relationship of several of the variables used in these studies to our "finalized" set of variables we have chosen.

all significantly related to bank failure. Similarly, using a logistic model, Cole and White (2012) find that the CAMELS components (equity to assets, non-performing assets to assets, ROA, cash to assets, investment securities to assets, and brokered deposits to assets) are all significantly related to bank failures in 2009. Ng and Roychowdhury (2014) examine the impact of using loan-loss reserves as "add-back" regulatory capital on bank failure during the 2007–2010 period. Counter to the assumption that higher capital leads to less chance of bankruptcy, they find a positive correlation between Loan Loss Reserves used as regulatory capital and the probability of bank failure.

Almanidis and Sickles (2012) use a mixture hazard model to predict bank failures over the 2008–2010 period. They produce an in-sample overall success rate of 94% for predicting failures in 2008-2009. Cole and Wu (2014) suggest that while hazard models seem to out-perform probit models, when information availability is accounted for, hazard models under-perform. Mayes and Stremmel (2014) arrive at a similar conclusion when they examine U.S. bank failures over the 1992–2012 period. They find that hazard models outperform logistic models in-sample with 97% success versus 80%. However, when they try to make out-of-sample forecasts, the logistic models perform better with an 86% success rate, which is lower than our out-of-sample success. They also note that prediction success drops significantly when trying to forecast two, three or four periods in advance. This is consistent with our observation that the use of timely information is critical.

The paper most related to ours is Jordan et al. (2010) in the sense that they use quarterly data and discriminant analysis to predict the failure of U.S. banks over a similar time period (i.e., 2007–2010). Our model produces superior results to theirs, as they have much lower prediction success rates ranging from 70.2% (four years in advance) to 77.6% (one year in advance), versus our success rates in the 90% range. This can be attributed to several factors. First, they use different variables to measure various bank characteristics. Second, our results show that using more recent (and available) data (up to one quarter in advance) leads to much higher prediction success rates, whereas they use data is that lagged one, two, three and four years. Logically the use of timely information would be of utmost importance during the most recent financial crisis as events were transpiring at a rapid pace. Third, their control sample only controls for state location and not bank size (as we do), hence we in all-likelihood have a better control sample for estimation purposes. Finally, Jordan et al. use a strict *Z*-score cut-off point to implement their predictions, whereas we also adopt the use of a "zone of ignorance" as prescribed by Altman (1968) that improves our results.

Several studies have also focused on the impact of non-traditional or off B/S activities and their impact on bank failures. Torna (2010) distinguishes between factors that cause healthy banks to become troubled and those that cause troubled banks to fail. He finds that larger "modern banking activities" (defined as investment banking and venture capital practices) increase the probability of banks becoming troubled, but that some of these activities (specifically brokerage services) reduced the likelihood of a troubled bank actually failing. DeYoung and Torna (2013) test whether the level of non-traditional banking activities (such as securities brokerage, insurance sales, venture capital, investment banking, and asset securitization) impacted the likelihood of bank failure during the crisis. Using a logit model, their results suggest that fee-based non-traditional activities (such as securities brokerage and insurance) may actually reduce the probability of failure while asset-based activities (investment banking and securitization) increased the probability of failure. De Jonghe (2010) finds complementary results for European banks over the 1992–2007 period using a different approach. He estimates that the "tail risk," which is defined as the bank's stock price sensitivity to a banking crisis, is higher for banks that engage in a higher proportion of non-traditional banking activities. While none of these studies attempt to forecast or predict failures, the results of these studies strengthen our decision to include a measure of off-balance sheet items in our analysis.

As the discussion above illustrates, there have been numerous studies of U.S. bank failures. Of course, one of the contributing factors to this is data availability – with over 1000 such failures over the 1985–93 period, and over 500 during 2008–2013. This is to be expected, given the large number of banks that operate in the U.S. versus other countries. This is a function of the evolution of the U.S. banking industry based on previously existing state banking restrictions. In contrast, as pointed out by Betz et al. (2014) "there are only a few dealing with European banks." They attribute this to "data limitations arising from relatively few direct bank failures in core Europe." The reduced number of failures available for examination is limited not only by the smaller "numbers" of European banks relative to U.S. banks, but also because government interventions (i.e., capital injections and/or forms of asset relief such as guarantees or asset protection) are more common in Europe. In contrast, in the U.S., the norm is to have failed banks purchased by healthy banks and continue operations. The net effect is that bank failures are much easier to identify in the U.S., which has provided easier access to sample evidence.

In order to deal with the difficulties in identifying a true bank distress situation for European banks, Betz et al. (2014) examine banks that undergo bankruptcy or liquidation, but also include those that receive government interventions or participate in a "distressed merger." Cihák and Poghosyan (2014) is an example of another European study that overcomes this data availability issue – in their case by examining financial media for words related to distress and investigating further. Similar to most of the U.S. studies discussed above, both of these studies use variables that are designed to capture CAMEL characteristics. It is interesting to note that despite the different institutional and legal environments, the European results are very similar to those of U.S. studies in terms of the signs and significance of the chosen variables, finding that both bank-specific variables and macro-financial factors are valuable in predicting bank distress. These similarities will become apparent in Section 4 when we compare our results to those of other studies. Finally, the predictive power of both of these studies is well below those of ours, similar to most of the U.S. studies discussed above.

3. Data and methodology

3.1. Data and sample creation

Our bank data is from two sources. The sample of failed banks is obtained from the failed bank list of the FDIC (http://www.fdic.gov/bank/individual/failed/banklist.html). We use all banks that failed between 2002 and 2009 for our "in-sample" analysis. There were 185 bank failures during this time period with the majority failing in 2009. We then examined 191 failed banks for which we could find data in the subsequent January 1, 2010–September 30, 2011 period to explore the "out-of-sample" predictive ability of our model.

Financial data for all banks (both failed and non-failed) was gathered from the Bankscope database. Of the 185 failed banks between 2002 and 2009, financial data for the quarter prior to failure was available for 132 banks. A matched sample of non-failed banks was created to use as a control sample. For each failed bank, a matched non-failed bank was selected based on geography (it had to be from the same state) and size (the closest in total assets.)

3.2. Variables

The variables chosen for our model were based on an extensive review of previous empirical studies related to bankruptcy in general, and to bank failures in particular. The seminal study for predicting firm bankruptcy was by Altman (1968), and this model, as well as variations thereof, is still widely used today. Altman used five factors to predict bankruptcy: (1) a liquidity measure (working capital/total assets); (2) a sustainable profitability measure (retained earnings/total assets); (3) an operating efficiency measure (EBIT/total assets); (4) a leverage factor (market value of equity/book value of total liabilities); and, (5) an asset turnover measure (sales/total assets).

In order to account for the distinct nature of financial institutions, several of the Altman (1968) factors need to be modified slightly. As a result, our model uses the Altman retained earnings to total assets variable, along with three other factors, which have been used in many previous studies related to bank distress[6]:

(1) liquidity measure = cash/total assets (Cash/TA).
(2) sustainable profitability measure = retained earnings/total assets (RE/TA).
(3) operating efficiency measure = return on average assets (ROA).
(4) leverage measure = equity/total assets (EQ/TA).

In addition to the Altman variables, which measure general financial health, we include bank-specific measures related to the following key areas of their success and stability: (1) reliance on loans; (2) loan quality; (3) capital adequacy; and, (4) off-balance sheet items. These factors represent a simplified version of the SCOR system used by the FDIC to assess bank stability, as described in Collier et al. (2005).

Brown and Serdar (2011) indicate that because loans are illiquid and deposits are liquid, banks that have a high proportion of loans are more likely to fail. We measure reliance on loans in three ways: first by the proportion of loans in the total asset mix, and then by the proportion of loans financed with deposits and either short-term (ST) funding or total borrowings. Specifically, the following three variables are used:

Loan1) Net Loans/Total Assets.
Loan2) Net Loans/(Total Deposits & Short-Term Funding).
Loan3) Net Loans/(Total Deposits and Borrowings).

All three of these measures are expected to be negatively related to the bank's overall financial health and variations of them have been used in previous studies. For example, both De Jonghe (2010) and Torna (2010) used Loans/Total Assets, while Almanidis and Sickles (2012) used Loans/Deposits.

Bank asset quality is a standard factor used in bank failure prediction models (for instance, see Jesswein (2009), Wheelock and Wilson (2000), Thomson (1991), and Martin (1977)). We measure bank loan quality using three ratios that have been used widely in previous studies[7]. All of these ratios are related to bank asset quality difficulties and would be expected to be positive related to the probability of default (and thus negatively related to the bank's health):

LoanQual1) Impaired Loans/Total Loans.
LoanQual2) Loan Loss Reserves/Total Loans.
LoanQual3) Net Charge Offs/Total Loans.

Estrella et al. (2000) find a negative relationship between bank capital and bank failure. In addition to using EQ/TA as discussed above, we also measure bank capital using two traditional Basel (risk-based capital) ratios, as have many other studies[8]:

Cap1) Tier 1 capital ratio = tier 1 capital/risk-weighted assets.
Cap2) Total capital ratio = total capital/risk-weighted assets.

A greater ratio for these measures indicates greater amounts of bank capital and thus we would predict would be positively related to the bank's health.

Cole and Wu (2009) suggest that the growing amount of off-balance sheet activities that banks are engaging in could have an impact on the probability of failure. Also, as discussed in Section 2, DeYoung and Torna (2013) find that non-traditional activities influence bank health. While there is certainly anecdotal evidence to this point (e.g., Lehman Brothers), it is not clear if this is true more generally. To address this issue, we include an off-balance sheet item variable in our model[9]. Off-balance sheet items include acceptances, documentary credits, loan guarantees, contingent liabilities, and other off-balance sheet items (including many of the "toxic" structured products that contributed to the downfall of many financial institutions). These are scaled by total assets.

3.3. Sample summary statistics

Table 1 provides summary statistics for our total sample of banks, as well as for the total matched sample of failed and matched banks[10]. The statistics are similar for both the total sample and the matched sample; however, the ratios are slightly weaker for the matched sample[11]. This is not surprising, since the matched sample was matched to failed banks according to size and location, and the financial crisis hit some regions harder than others, as discussed in Aubuchon and Wheelock (2010). A comparison of the summary statistics for the sample of failed banks to the matched sample provides overall results that one would expect for the most part. As expected, and consistent with almost previous empirical evidence, the failed banks have much lower values for RE/TA, EQ/TA, ROA, Capital Adequacy ratios, and higher ratios of Impaired Loans/Total Loans (LoanQual1), Loan Loss Reserves/Total Loans (LoanQual2) and Net Charge-Offs/Total Loans (LoanQual3). The ratios are similar for both groups for the Loan Reliance measures (Loan1/2/3), which differs from the findings of Torna (2010) and Almanidis and Sickles (2012) that troubled banks had higher ratios.

The reported Cash/TA ratios in Table 1 are actually higher for the failed banks, contrary to what one might initially expect. This contradicts the summary evidence of Almanidis and Sickles (2012), Cole and White (2012), DeYoung and Torna (2013), and Torna (2010). It is interesting to note however, that in these papers,

---

[6] For example, Almanidis and Sickles (2012) use both Cash/TA and ROA; Berger and Bouwman (2014) use EQ/TA, ROE and Cash/TA; Betz et al. (2014) use ROA; Cihák and Poghosyan (2014) use ROA and Liquid Assets/TA; both Cole and White (2012) and Cole and Wu (2014) use EQ/TA and ROA, while Cole and White (2012) also use (Cash + Cash Due)/TA; De Jonghe (2010) uses EQ/TA, ROE and Liquid Assets/TA; DeYoung and Torna (2013) use EQ/TA, ROA and Cash/(Borrowings plus Deposits); Ng and Roychowdhury (2014) use ROA and (Cash + Amounts due from other banks)/Deposits; and, Torna (2010) uses ROA and Cash/(Borrowings + Deposits).

[7] For example, Almanidis and Sickles (2012) include all three of our measures. Cihák and Poghosyan (2014), Jin et al. (2011), and Ng and Roychowdhury (2014) use both Impaired Loans/TA and Loan Loss Reserves/Total Assets, while Betz et al. (2014) use Loan Loss Reserves/Total Assets. Cole and White (2012), Cole and Wu (2014), DeYoung and Torna (2013), and Torna (2010) all use Non-Performing Loans/Total Assets.

[8] For example, Berger and Bouwman (2014), and Betz et al. (2014) use both, while Almanidis and Sickles (2012), Jin et al. (2011), and Ng and Roychowdhury (2014) use the Tier 1 ratio.

[9] We attempted to use various off-balance sheet measures; however, this measure was the only one with consistent data across the firms in our sample due to data limitations in Bankscope.

[10] All statistics represent figures for the prior quarter. Refer to the Appendix for a complete description of the variables.

[11] Table 1 also includes summary statistics for the Z-score, which will be discussed later.

Table 1
Summary statistics. Failed versus matched banks (2002–2009).

Variable	Total sample		Failed Banks		Matched SAMPLE	
All as percentages, except for *Z* scores	Mean	Median	Mean	Median	Mean	Median
Cash/TA	10.05	7.55	11.47	9.67	9.03	6.76
RE/TA	4.40	4.57	−7.18	−6.31	2.92	2.95
EQ/TA	11.67	9.93	2.46	2.36	9.73	9.20
ROA	0.00	0.01	−11.36	−8.79	−0.47	0.00
Loan1 (loans/TA)	63.72	67.01	68.15	70.40	60.64	68.69
Loan2 (loans/(deposits + ST funding))	74.70	77.94	73.50	75.90	71.28	77.37
Loan3 (loans/(deposits + borrowing))	72.20	75.48	70.70	72.12	68.07	75.32
LoanQual1 (impaired loans/loans)	2.71	1.48	15.55	13.18	3.79	2.13
LoanQual2 (LLR/loans)	1.62	1.39	4.10	4.13	1.74	1.48
LoanQual3 (net charge-offs/loans)	0.53	0.15	3.22	2.03	5.15	0.18
Cap1 (tier 1 cap ratio)	17.02	13.00	2.75	2.50	12.30	11.30
Cap2 (total cap ratio)	18.16	14.20	3.75	3.80	13.37	12.50
Offbs (OffBS/TA)	15.49	9.41	12.04	7.00	19.33	11.91
Z-score	1.55	1.41	−1.17	−1.13	1.28	1.31

Table 2
Correlation coefficients (Matched and Failed Sample – 2002–2009).

	Cash/TA	RE/TA	EQ/TA	ROA	Loan1	Loan2	Loan3	LoanQual1	LoanQual2	LoanQual3
Cash/TA	1.00									
RE/TA	-0.21^{**}	1.00								
EQ/TA	0.21^{**}	0.51^{**}	1.00							
ROA	−0.11	0.65^{**}	0.51^{**}	1.00						
Loan1	-0.52^{**}	−0.07	0.02	-0.18^{**}	1.00					
Loan2	-0.52^{**}	0.13^{*}	0.20^{**}	−0.05	0.96^{**}	1.00				
Loan3	-0.51^{**}	0.11	0.21^{**}	−0.06	0.97^{**}	0.98^{**}	1.00			
LoanQual1	0.02	-0.54^{**}	-0.43^{**}	-0.49^{**}	0.22^{**}	0.10	0.11	1.00		
LoanQual2	−0.02	-0.52^{**}	-0.35^{**}	-0.55^{**}	0.34^{**}	0.24^{**}	0.25^{**}	0.65^{**}	1.00	
LoanQual3	0.02	-0.57^{**}	-0.43^{**}	-0.66^{**}	0.17^{**}	0.05	0.07	0.58^{**}	0.47^{**}	1.00
Cap1	0.52^{**}	0.19^{**}	0.78^{**}	0.27^{**}	-0.19^{**}	−0.11	−0.11	-0.24^{**}	-0.24^{**}	-0.23^{**}
Cap2	0.51^{**}	0.20^{**}	0.78^{**}	0.27^{**}	-0.18^{**}	−0.09	−0.09	-0.25^{**}	-0.24^{**}	-0.24^{**}
Offbs	−0.12	0.13^{*}	0.25^{**}	0.09	0.20^{**}	0.34^{**}	0.38^{**}	−0.11	0.12	−0.01

	Cap1	Cap2	Offbs
Cap1	1.00		
Cap2	1.00^{**}	1.00	
Offbs	0.06	0.07	1.00

* Significant at 5% level.
** Significant at 1% level.

Almanidis and Sickles (2012) find that Cash/TA is significantly and positively related to distress in their distress prediction model (as do we), while the other three studies find it is insignificant in their respective models. Almanidis and Sickles (2012) suggest that "One explanation of this could be, after controlling for profitability, that banks that remain cash idle have a higher opportunity cost. It would only stand to reason for these banks to be costly and inefficient." This argument is consistent with those of Boot and Thakor (2010), Berger and Bouwman (2009), and Berger et al. (2014) who all suggest that banks that find themselves in trouble and receive liquidity or guarantees may not use such vehicles to provide loans, but rather to build up their own liquidity. In other words, this could be due to the fact that banks that find themselves in difficult situations, whether they receive support or not, are reluctant to issue new loans and instead build up liquidity in anticipation of future troubles.

The theory outlined in the paragraph above is also consistent with the lower values for Loans/TA observed for failed banks in Table 1. Both of these patterns are expected when firms find themselves in trouble, and it is interesting to note that the cash pattern is strong when we use timely quarterly data, but when we used less timely annual data the pattern was much weaker. In particular, when we examined annual data we found that failed firms had slightly higher mean cash ratios than the matched sample (2.9% versus 2.4%), but lower median ratios (1.8% versus 2.1%). Also, when annual data was used, we found that failed firms had higher levels of Loans/TA, contrary to the pattern in Table 1 based on quarterly data. Finally, Berger and Bouwman (2009) also provide empirical evidence supporting their argument, noting that firms with higher lagged capital ratios will hold "fewer liquid assets," which is consistent with the results we provide in Table 1.

Table 2 includes correlation coefficients for the variables used in this study. Overall there are no surprises in the table with variables measuring similar characteristics displaying high correlations as one would expect. For example we observe high correlations across the various variables used to measure Loan Reliance (Loan1/2/3), Loan Quality (LoanQual1/2/3), and Capital Adequacy (Cap1 and Cap2). We also observe high correlations of 0.65 between ROA and RE/TA, and of 0.78 between EQ/TA and the Tier 1 Capital and Total Capital ratios. The later observation implies that we will have to make a decision as to whether to use either EQ/TA or the Capital ratios in our model. Finally, it is interesting to note the high negative correlations between our three loan quality measures with both EQ/TA and RE/TA, and to a lesser extent with the capital ratios. These observations are consistent with the summary statistics reported in Table 1 which suggest that failed banks have both lower equity values and weaker loan quality than surviving banks. This is consistent with the argument of Mehran and Thakor (2011) who suggest that maintaining capital reserves provides a secondary indirect benefit to banks, aside from the obvious direct

Table 3
Discriminant analysis results.

Variable	Predicted sign	Z-score
Panel A coefficients		
Constant	?	0.048
Cash/TA	+	−1.769*
RE/TA	+	8.008**
EQ/TA	+	12.856**
ROA	+	0.560**
Loan1 (loans/TA)	− (?)	0.286
LoanQual2 (LLR/loans)	−	−21.285**
Offbs (Off BS/TA)	−	−0.050*
Success Rate		91.7%

Variable	Z-score
Panel B (relative contribution factors)	
Cash/TA	17.89
RE/TA	102.11
EQ/TA	188.51
ROA	5.60
Loan1 (loans/TA)	2.10
LoanQual2 (LLR/loans)	1030.22
Offbs (Off BS/TA)	0.08

Model (Z-score)	Predicted group		
	Failed	Not Failed	Total
Panel C (classification summary)			
Bankrupt	117 (88.6%)	15 (11.4%)	132
Not Bankrupt	6 (5.0%)	115 (95.0%)	121

* Significant at 5% level.
** Significant at 1% level.

benefit of increasing survival probability. Namely, they argue that "greater loan monitoring induced by higher capital enhances the value of the relationship loan portfolio." This greater monitoring could in turn contribute to the higher loan quality displayed by better capitalized firms, which captures two of the patterns we observe in Table 1 for surviving firms.

3.4. Multiple discriminant analysis (MDA)

The index we use to classify bank health is determined using MDA, similar to Altman's Z factor for predicting bankruptcy. Altman (1968) applied this statistical technique to his sample of 66 firms over the 1946–1965 period for the purpose of distinguishing firms that were likely to go bankrupt from those that were likely to avoid bankruptcy. During his sample period, 33 firms go bankrupt, while the other 33 are still in existence at the end of the period. Using MDA, he was able to predict with 95% accuracy, which firms would go bankrupt and which firms would not. Altman et al. (1977) are able to achieve similar success for 111 firms over the 1969–1975 period using a modified set of independent variables in the MDA specification.

MDA involves choosing mutually exclusive groups with regards to some qualitative trait (e.g., bankrupt versus non-bankrupt firms). The next step involves deriving a linear combination of characteristics that "best" discriminates between the two groups. The analysis considers an entire profile of characteristics common to the relevant firms, as well as the interaction of these properties, and transforms them into a univariate statistic, commonly referred to as a Z-score.

One of the main advantages of this technique is that it allows the analysis of the entire variable profile of a firm simultaneously, rather than sequentially examining the individual characteristics. This allows it to be used to efficiently examine and categorize a large number of firms. In our case it permits us to determine Z-scores for a large sample of banks that have not failed and that are not used in our matched sample. These banks can then easily be categorized by reference to their Z-scores as those that "fit the profile" of failed or healthy banks. Summary statistics that will be discussed in the next section confirm that in general, banks with higher Z-scores appear to be much healthier in terms of the variables discussed above. This combination of effectiveness and ease of implementation makes it a very functional model.

We examined a large number of combinations of the variables discussed above, and in many of the papers reviewed in the literature, in order to determine which version of the model would work best at predicting which firms would go bankrupt using MDA[12]. The following variables provided the highest in-sample success rate in predicting correctly which firms would fail and those that didn't (with 91.7% success) – Cash/TA, RE/A, EQ/TA, ROA, Loans/TA (i.e., Loan1), Loan Loss Reserves/Loans (i.e., LoanQual2), and Off-Balance Sheet/TA (i.e., Off-BS). The equation for this Z-score is given below:

$$Z1 = \beta_0 + \beta_1 Cash/TA + \beta_2 RE/TA + \beta_3 EQ/TA + \beta_4 ROA + \beta_5 Loans/TA + \beta_6 Loan\ Loss\ Reserves/Loans + \beta_7 Off - BS \quad (1)$$

Without going into too much detail on which variables and which combinations thereof worked better or worse, one decision is worthy of comment. As suspected, due to the high correlation between EQ/TA and the two risk-based Capital ratios, we had to drop two of the three variables. We chose to use the specification "without" the Capital Adequacy ratios (i.e., Cap1/2) since it worked better when EQ/TA was included and they were excluded. While we did so solely to improve our prediction accuracy, this result is consistent with the conclusions of several previous empirical studies such as Berger and Bouwman (2014), Demirgüç-Kunt and Huizinga (2010), and Mayes and Stremmel (2014). For example, Mayes and Stremmel (2014) suggest that "our findings suggest that the more complex a bank is, the more effective is the leverage ratio compared to the risk-based capital ratio." Berger and Bouwman (2014) also find better results using EQ/TA and suggest that they prefer to use EQ/TA because "regulatory ratios mix capital with credit risk, which is already controlled for in our regressions."

4. Results

This section discusses the results from discriminant analysis. As mentioned in the previous section, we examined numerous combinations of variables and now focus our discussion on the combination that provides the best predictive power. Univariate significance levels denoted in Panel A of Table 3 indicate that RE/TA, EQ/TA, ROA and Loan Loss Reserves/Loans are significant at the 1% level[13]. Cash/TA and Off-BS/TA are univariately significant at the 5% level, while Loans/TA is insignificant; however, all of the variables add some explanatory power, as evidenced by the improvement in predictive power that occurs when they are included. This is also evident if we look at the relative contribution factor of the variables[14]. These factors are reported in Panel B of Table 3.

Panel A of Table 3 presents the coefficients for each of the variables included in our model and most of the coefficients are of the

[12] For example, we used all four "Altman-like" variables and the Off-Balance Sheet measure in every possible combination with each of the various measures of Loan Reliance (Loan1/2/3), Loan Quality (LoanQual1/2/3) and Capital Adequacy (Cap1/2).

[13] The univariate significance levels are determined using F-tests that examine the individual discriminating ability of each variable by relating the difference between the average values of the ratios in each group to the variability (or spread) of values of the ratios within each group. The common F-value tests the null hypothesis that the observations come from the same population. If the null is rejected, then it makes sense to move forward and use the variable to try to discriminate between the two groups.

[14] This measure is constructed by multiplying the variable standard deviation by its coefficient in the discriminant function.

Table 4
Predicted group statistics (in-sample results – 2002–09).

Variable	PreGrp 0 (Fail)		PreGrp 1 (Healthy)	
All as percentages, except for *Z* scores	Mean	Median	Mean	Median
Cash/TA	14.79	11.81	9.87	7.41
RE/TA	−5.96	−5.70	4.80	4.80
EQ/TA	6.32	5.74	11.88	10.02
ROA	−5.69	−4.92	0.02	0.60
Loan1 (loans/TA)	65.55	68.71	63.65	66.92
Loan2 (Loans/(deposits + ST funding))	72.64	75.20	74.78	78.09
Loan3 (loans/(deposits + borrowing))	70.42	72.99	72.26	75.59
LoanQual1 (Impaired Loans/Loans)	10.53	10.46	2.41	1.39
LoanQual2 (LLR/loans)	3.74	3.68	1.54	1.37
LoanQual3 (net charge-offs/loans)	2.56	2.15	0.45	0.13
Cap1 (tier 1 cap ratio)	8.57	7.16	17.34	13.20
Cap2 (total cap ratio)	9.93	8.50	18.47	14.30
Offbs (OffBS/TA)	14.68	7.35	15.52	9.51
Z-score	−0.52	−0.39	1.63	1.44

predicted sign. Loan Loss Reserves/Loans has a large negative coefficient as expected, and provides the highest relative contribution to our model according to the factors presented in Panel B. This finding is consistent with those of previous studies with respect to loan quality variables in both U.S. and European studies; although, as discussed in Section 3 some used alternative measures such as non-performing loans or impaired assets, and several studies used more than one measure. We used only the one measure due to the high correlation observed among the three variables in Table 2, which resulted in enhanced predictive power.

As expected, EQ/TA and RE/TA have large positive coefficients, and represent our second and third most relevant factors, respectively, according to the relative contribution factors reported in Panel B. As mentioned previously, we did not find any previous bank failure studies that used RE/TA. On the other hand several previous studies used EQ/TA and our results are in line with the results of these studies including those of: Berger and Bouwman (2009), Cihák and Poghosyan (2014), Cole and White (2012), Cole and Wu (2014), De Jonghe (2010); and, Deyoung and Torna (2013). It is noteworthy that the first two studies listed above examine European bank failures, and find that EQ/TA is of similar importance as is found in U.S. studies. As noted in Section 3, the fact that EQ/TA worked better in our model than the risk-based capital ratios is consistent with the findings of several previous studies.

ROA has a much smaller but positive and significant coefficient as expected. The sign and lower impact (than capital) is consistent with the results of most previous studies such as Cole and White (2012) and Deyoung and Torna (2013). Berger and Bouwman (2014), and De Jonghe (2010) obtain similar using ROE. Finally, Cole and Wu (2014) find ROA is significant in their hazard model, but not in their logit model; while Almanidis and Sickles (2012) find it is significant in only two of the four models they examine. Off-BS/TA has a negative coefficient, as expected, however it is quite small; albeit significant at the 5% level. Unfortunately, this measure is very broad but was the only one with consistent data across the firms in our sample due to data limitations in Bankscope. The lack of effectiveness of this measure highlights the fact that using such an aggregate measure of off-balance sheet items is not as relevant as is an assessment of the "quality" and associated "underlying risks" associated with these items; although that is not the primary focus of our study.

Loans/TA has a positive coefficient, contrary to expectations, but it is small and insignificant. This contradicts the result of Torna (2010) who finds Loans/TA are positively related to failure, but is consistent with the finding of Almanidis and Sickles (2012) that the Loans-to-Deposits ratio is negatively related to failure. Finally, contrary to what we initially expected, the Cash/TA variable has a negative coefficient that is significant at the 5% level. Previous empirical evidence on this issue is mixed. Ng and Roychowdhury (2014) found that liquidity was negatively and significantly related to bank failure, while Almanidis and Sickles (2012) found it was positively and significantly related to failure, as do we. Six other studies found liquidity to be insignificant (e.g., Cole and White (2012), Berger and Bouwman (2014), De Jonghe (2010), DeYoung and Torna (2013), Torna (2010), and Cihák and Poghosyan (2014)). Our finding is consistent with the summary statistics we reported in Table 1, as well as the arguments of Boot and Thakor (2010), Berger and Bouwman (2009), and Berger et al., 2014 that were advanced earlier. These studies suggest that banks that are in danger of severe distress may be building up liquidity and reducing the amount of loans issued in anticipation of future troubles[15]. Our evidence seems consistent with this argument; although Cash/TA is not one of our most important factors in predicting distress.

Panel C of Table 3 provides the classification results for our model using quarterly data – with a 91.7% success rate in classifying the banks. This exceeded the 85.3% success rate we achieved when we used annual rather than quarterly data. This suggests the importance of using more timely information, which becomes very critical when banks are facing financial distress. Our in-sample success rate is better than that of most previous studies, including Jordan et al. (2010) who produce a success rate of 86.8% also using a discriminant model and a similar sample[16]. However, the real test of the usefulness of our model lies in its ability to distinguish between healthy banks and those that are "troubled" and therefore run higher risk of failure, as discussed below.

We examine the model's effectiveness by applying it to our entire sample (i.e., including all banks – beyond "just" the failed and matched sample groups) of U.S. banks using previous quarter-end data. Table 4 presents summary statistics for two groups that were formed based on *Z*-scores calculated using previous quarter-end data and applying the coefficients from the model above. The groups are labeled PreGrp0 (those that our models predict will fail) and PreGrp1 (those that our model predicts are healthy and will not fail).

Overall, the summary statistics are very impressive from a classification perspective. The banks classified as facing a high risk of

[15] Consistent with this argument, while Torna (2010) finds liquidity is insignificant in his overall regression, he finds it is positively and significantly related to failure for banks that were "pre-classified" as troubled.

[16] Our in-sample success is slightly below that achieved by Alamanidis and Sickles (2012), and Mayes and Stremmel (2014), both of which use hazard models.

Table 5
Predicted group statistics (out-of-sample results – Q1 2010–Q3 2011).

Variable	PreGrp 0 (Fail)		PreGrp 1 (Healthy)		Actual failed firms	
All as percentages, except for *Z* scores	Mean	Median	Mean	Median	Mean	Median
Cash/TA	15.29	12.94	10.92	8.44	11.92	10.83
RE/TA	−6.55	−6.26	5.00	5.08	−5.82	−5.73
EQ/TA	6.74	6.35	11.78	10.18	4.31	3.81
ROA	−2.68	−1.81	0.57	0.71	−5.04	−5.02
Loan1 (loans/TA)	63.36	66.46	61.06	63.86	69.28	70.10
Loan2 (loans/(deposits + ST funding))	70.89	73.14	72.28	74.43	74.30	74.64
Loan3 (loans/(deposits + borrowing))	68.65	71.23	70.13	72.40	71.84	72.27
LoanQual1 (impaired loans/loans)	11.26	10.65	3.13	1.86	16.06	16.97
LoanQual2 (LLR/loans)	3.90	3.74	1.78	1.53	3.92	3.98
LoanQual3 (net charge-offs/loans)	1.82	1.26	0.48	0.14	2.73	2.42
Cap1 (tier 1 cap ratio)	10.19	8.55	17.69	14.03	4.95	4.37
Cap2 (total cap ratio)	11.47	9.92	18.87	15.25	6.58	6.10
Offbs (OffBS/TA)	12.86	6.45	15.38	8.89	8.39	5.03
Z-score	−0.55	−0.39	1.56	1.40	−0.74	−0.78
Number of observations	2,759 (all 7 quarters) (average 394.1 per quarter)		45,869 (all 7 quarters) (average 6,552.7 per quarter)		191 (all 7 quarters) (average 27.3 per quarter)	

Table 6
Out-of-sample predictions for failed banks.

	Predicted group		
	Fail	Not fail	Zone of ignorance
Using cut-off of 0.05	171 (89.5%)	20 (10.5%)	NA
Using cut-off to 90 zone (.05–.1408)	171 (89.5%)	17 (8.9%)	3 (1.6%)
Using cut-off to 95 zone (.05–.4587)	171 (89.5%)	8 (4.2%)	12 (6.3%)

failure (i.e., PreGrp0 banks) have very weak balance sheets with negative average and median RE/TA ratios and very low EQ/TA values (average 6.3% and median 5.7%), and are unprofitable (negative average and median ROAs). They possess a higher percentage of poorly performing loans, with high percentages of Impaired Loans, Loan Loss Reserves and Net Charge-Offs relative to Total Loans – averages of 10.5%, 3.7% and 2.6%. They also possess low Capital Adequacy ratios with average Tier 1 and Total Capital ratios of 8.6% and 9.9%, relative to 17.3% and 18.5% for the banks classified as being healthy. It is noteworthy that the average Total Capital ratios for those that our model predicted would fail are slightly below the Basle II and III requirements of 10% and 10.5%, respectively, while the median value of 8.5% is well below these requirements. Similarly, the Tier 1 average ratio is only slightly above the Basle III requirement of 8.5%, while the median value of 7.2% is well below this requirement. In contrast, the banks classified as healthy possess average and median capital ratios that are much higher than the Basle II and III requirements. This type of effectiveness, combined with the ease of implementation (i.e., we were able to easily classify over 6000 banks per quarter using our *Z*-score formula), makes this a very functional model.

5. Out-of sample predictions

Jones (1987) pointed out the importance of using a "hold-out sample" to test the external validity of bankruptcy models, while Bellovary et al. (2007) found that less than half of the studies they reviewed did so. However, we believe this is an important consideration, given the potential applications of our model, as discussed above. In order to conduct our "out-of-sample" analysis, we use data up to the end of 2009 for our "in-sample" analysis discussed above. We then use "out-of-sample" observations for bankrupt U.S. banks between January 1, 2010 and September 30, 2011, which provided us with a sample of 191 failed banks. This differs from Jordan et al. (2010) for example, who randomly select half of their sample as the "training sample" and test it against the other half as the "hold-out" sample. All banks are classified using the discriminant model described in Table 3, using quarterly data that is "lagged" one quarter to ensure data availability for use in our model. We argue that by estimating our model using ex ante quarterly data and then applying it to firms in a subsequent quarter using information that is publicly available, our model is superior from an economic point of view for predicting future financial distress.

Table 5 provides the summary statistics for three groups – those that our model predicted would fail (PreGrp0), those our model predicted would not fail (PreGrp1), and those that actually failed during the seven-quarter period. The results in Table 5 confirm the relevance of our model in an out-of-sample setting, as we observe that traditional ratios are much weaker for those in PreGrp0. In particular, we see that these firms have much smaller (and negative) average and median values for RE/TA and ROA, much smaller values for EQ/TA, Tier 1 Capital and Total Capital, and larger values for all three measures which are inversely related to loan quality. The firms that actually went failed over this period had even weaker ratios on average than the PreGrp0 firms with respect to loan quality and capital. In fact, the average (median) capital ratios for Tier 1 Capital of 5.0 (4.4) and Total Capital of 6.6 (6.1) were well below the Basle II and III capital guidelines. Again, these statistics point to the importance of maintaining capital adequacy and loan quality.

Table 6 provides the classification summary results of applying our model using lagged quarterly data. The top row shows the prediction success rate of 89.5% using a strict "cut-off" point (which was a *Z*-score of 0.05). It should be noted that the use of quarterly data led to a dramatic improvement in predictive power over the use of annual data, which was only 55% successful "out-of-sample". This reiterates the importance of using timely data, which was noted previously when referring to "in-sample" model performance. However, the "out-of-sample" model performance is much more dramatically improved by the use of quarterly data.

Table 7
Z-statistics (out-of-sample).

Model/Variable					
Z	# Observations	Mean	Median	Max	Min
Panel A: for "Missed" failed firms					
Using .05 cut-off	20	0.54	0.39	1.38	0.07
10/90 Band (0.05–0.1408)	17	0.62	0.40	1.38	0.16
5/95 Band (0.05–0.4587)	8	1.01	1.07	1.38	0.48
Panel B: for "correctly predicted" failed firms (out-of sample)					
Using .05 cut-off	171	−0.89	−0.87	0.03	−2.11
10/90 Band (0.05–0.1408)	174	−0.88	−0.87	0.14	−2.11
5/95 Band (0.05–0.4587)	182	−0.82	−0.82	0.40	−2.11

The quarterly model successfully predicted 171 (89.5%) of the 191 banks in our sample that failed between January 1, 2010 and September 30, 2011. These results are impressive and they far exceed the out-sample success in predicting failed banks of 35% reported by Thomson (1991) in his study of U.S. bank failures during the 1980s. They are also well above the 27% success rate in predicting downgrades using the SCOR system that is reported by Collier et al. compared to more recent studies, of those that conduct out-of-sample predictions, we outperform the success rates of the best models of Cihák and Poghosyan (2014) (68%), Jordan et al. (2010) (78%), Mayes and Stremmel (2014) (86%), and Betz et al. (2014) (89%). While prediction success is not the sole purpose of constructing our model, our success rate does indicate it is effective.

Further examination of the "misclassified" (as healthy) firms suggests that our model was not that far off the mark with regards to these firms. For example, Panel A of Table 7 provides summary statistics for the "misclassified" banks. These statistics show that while these banks did not have *Z*-scores below the cut-off value of 0.05, they had an average of 0.54 and a median of 0.39, with a maximum of 1.38. We can relate this result to Altman's reference to a "zone of ignorance" of *Z*-scores where "misclassifications can be observed." In his study, he indicated that 10 of 14 misclassifications occurred for *Z*-scores between 1.81 and 2.67.

We construct two such "zones" of ignorance for our model's predictions. These can be viewed as ranges of *Z*-scores where the firm is not clearly at the highest risk of failure, but they are at substantial risk and are thus worthy of additional scrutiny. We construct our first zone, which we label the 5/95 zone by making reference to our original failed and matched bank sample and using the 95th percentile *Z*-score for the failed banks (0.4587) and the lower of the 5th percentile for the non-failed group (0.0891) and the "cut-off" point (0.05). Hence the 5/95 zone of ignorance includes *Z*-scores between 0.05 and 0.4587, and banks with *Z*-scores in this region would be classified as lying in a "gray area." This is a fairly wide zone, so we also constructed a more stringent zone, denoted as the cut-off to 90 zone, using the 90th percentile for the failed banks of the in-sample banks (0.1408) and the lower of the 10th percentile for the non-failed group (0.4468) or the model "cut-off" point (0.05). Hence the cut-off to 90 zone would be from 0.05 to 0.1408.

The 2nd and 3rd rows in Table 6 provide the results using these two zones of ignorance. When we use the wider 5/95 zone, all but 8 (4.2%) of the firms that failed would have fallen within this zone. When the more narrowly defined zone (cut-off to 90) is used, only 17 (8.9%) would have been classified as "healthy." Panel A of Table 7 shows that the failed firms that were mis-classified as healthy using the 10/90 (5/95) bands had average *Z*-scores of 0.62 (1.01), with medians of 0.40 (1.07), so were generally not given high values for "health." Finally, Panel B of Table 7 reports summary statistics for banks that were classified as likely to fail. The mean (median) *Z*-scores using the strict cut-off, the 10/90 zone and the 5/95 zone were −0.89 (−0.87), −0.88 (−0.87) and −0.82 (−0.82), respectively, while the maximum values were 0.03, 0.14, and 0.40 – all of which suggests they were not classified as being very healthy according to their *Z*-scores.

Combining the results of Tables 6 and 7, we can say that our model works remarkably well out-of-sample in terms of the success rate in predicting which firms would fail, and also in the sense that they did not provide a false sense of security regarding the health of any of these failed banks. This type of "out-of-sample" success illustrates why our model represents such an important contribution. However, in addition to the out-of-sample prediction success, Table 5 showed that our model clearly differentiates between large numbers of healthy and distressed firms according to traditional ratios. Such an effective and easy to implement model will be of obvious interest to banks' internal risk management, to regulators, as well as to others with a direct interest in assessing bank financial health such as analysts, fund managers, and financial transaction counter-parties.

6. Conclusions

We examine the failures of 132 U.S. banks over the 2002–2009 period using MDA, after constructing a matched sample of non-failed banks for comparison purposes. Our model successfully distinguishes between banks that would fail and those that wouldn't 92% of the time using our in-sample quarterly data. Our two most important variables are related to bank capital and loan quality, as one might expect; although bank profitability is also important. These results confirm previous U.S. and international empirical evidence, as well as previous theoretical arguments, as discussed herein. Interestingly we also find that cash holdings are *positively* related to probability of distress. While this result initially seems counter-intuitive, previous empirical evidence regarding the influence of the liquidity variable is mixed and inconclusive, and our result is consistent with some of it. Our finding is also consistent with established theoretical arguments that suggest that troubled banks may build up liquidity at the expense of extending loans. We believe that our study is more likely to pick up this effect than many previous studies since we utilize more timely information. Even more important than our in-sample prediction success is the fact that our results demonstrate that our model can easily be applied to a large number of firms and do an excellent job of distinguishing healthy banks from those that are at high risk of failure (even those that do not fail).

We then apply our model out-of-sample to examine the failure of 191 banks during 2010–11, and observe that our out-of-sample predictions improve dramatically when we used quarterly rather than annual data. This highlights the importance of using timely information and regular monitoring banks, particularly those that appear to be troubled. Our out-of-sample predictive accuracy falls in the 90–95% range. This out-of-sample prediction success is impressive, as is the out-of-sample ability of our model to differentiate large numbers of healthy versus troubled banks using available information. Combining this effectiveness with its ease of implementation makes this a very attractive and functional model that will be of obvious interest to banks' internal risk management, to regulators, as well as to others with a direct interest in assessing bank financial health such as analysts, fund managers, and financial transaction counter-parties. For example, it could easily be used to assist the FDIC in preparing which banks should be on their "watch list" of troubled banks, which is used to prioritize site visits, among other things. It can also be used to assess the overall health of the banking system.

Appendix A. Variable descriptions

Cash/TA: CASH/TA (data2075/data2025).
RE/TA: RE/TA (data6320/data2025).
EQ/TA: Equity/TA (data4009).
ROA: REV/TA (data4024).
Loan1 (Loans/TA): Reliance on Loan (data4032)
Loan2: Net Loans/Deposits & ST Funding (data4033).
Loan3: Net Loans/Total Deposits and Borrowing (data4034).
LoanQual1: Impaired Loans/Total Loans (data4004).
LoanQual2: Loan Loss Reserves/Total Loans (data4001).
LoanQual3: Net Charge Offs/Total Loans (data4005).
Cap1: Tier 1 ratio (data4007).
Cap2: Total capital ratio (data4008).
Offbs: Off Balance Sheet Items/TA (data2065/data2025).

References

Almanidis, P., Sickles, R., 2012. Banking Crises, Early Warning Models, and Efficiency. Working paper.
Altman, E.I., 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. Journal of Finance, 589–609.
Altman, E.I., Haldeman, R.G., Narayanan, P., 1977. Zeta analysis, a new model to identify bankruptcy risk of corporations. Journal of Banking and Finance, 29–54.
Aubuchon, C., Wheelock, D., 2010. The geographic distribution and characteristics of U.S. bank failures, 2007–2010: do bank failures still reflect local economic conditions? Federal Reserve Bank of St. Louis Review, 395–415.
Bellovary, J.L., Giacomino, D.E., Akers, M.D., 2007. A review of bankruptcy prediction studies: 1930 to present. Journal of Financial Education 2007, 1–42.
Berger, A., Bouwman, C., 2009. Bank liquidity creation. The Review of Financial Studies 22 (9), 3779–3837.
Berger, A., Bouwman, C., 2014. How does capital affect bank performance during financial crises? Journal of Financial Economics 109 (1), 146–176.
Berger, A., Bouwman, C., Kick, T., Schaeck, K., 2014. Bank Risk Taking and Liquidity Creation Following Regulatory Interventions and Capital Support. Working paper.
Betz, F., Oprica, S., Peltonen, T., Sarlin, P., 2014. Predicting distress in European banks. Journal of Banking and Finance 45, 225–241.
Bhattacharya, S., Thakor, A.V., 1993. Contemporary banking theory. Journal of Financial Intermediation 3, 2–50.
Boot, A., Thakor, A.V., 2010. The accelerating integration of banks and markets and its implications for regulation. In: Berger, A., Molyneux, P., Wilson, J. (Eds.), The Oxford Handbook of Banking, pp. 58–90.
Brown, C., Serdar, D., 2011. Too Many to Fail? Evidence of Regulatory Forbearance in Bank Failures when the Banking Sector is Weak. Review of Financial Studies 24 (4), 1378–1405.
Cihák, M., Poghosyan, T., 2014. Anatomy of distress in European banks before and during the crisis. Retrieved from <http://www.voxeu.org/article/anatomy-distress-new-insight-probability-bank-turmoil>.
Cole, R., White, L., 2012. Déjà vu all over again: the causes of U.S. commercial bank failures this time around. Journal of Financial Services Research 42, 5–29.
Cole, R., Wu, Q., 2014. Hazard Versus Probit in Predicting U.S. Bank Failures: A Regulatory Perspective Over Two Crises. Working paper.
Cole, R., Gunther, J., 1998. Predicting bank failures: a comparison of on- and off-site monitoring systems. Journal of Financial Services Research 13, 103–117.
Cole R., Wu, Q., 2009. Predicting Bank Failures Using a Simple Dynamic Hazard Model. Working paper.
Collier, C., Forbush, S., Nuxoll, D., O'Keefe, J., 2005. The SCOR system of off-site monitoring: its objectives, functioning, and performance. FDIC Banking Review 15 (3), 17–32.
De Jonghe, O., 2010. Back to the basics in banking? A micro-analysis of banking system stability. Journal of Financial Intermediation 19 (3), 387–417.
Demirgüç-Kunt, A., Huizinga, H., 2010. Bank activity and funding strategies: The impact on risk and returns. Journal of Financial Economics 98 (3), 626–650.
Deyoung, R., Torna, G., 2013. Nontraditional banking activities and bank failures during the financial crisis. Journal of Financial Intermediation 22, 397–421.
Diamond, D., 1984. Financial intermediation and delegated monitoring. Review of Economic Studies 51 (3), 393–414.
Diamond, D., Dybvig, P., 1983. Bank runs, deposit insurance, and liquidity. The Journal of Political Economy 91 (3), 401–419.
Espahbodi, P., 1991. Identification of problem banks and binary choice models. Journal of Banking and Finance 15 (1), 53–71.
Estrella, A., Park, S., Peristiani, S., 2000. Capital Ratios as Predictors of Bank Failure. FRBNY Economic Policy Review, pp. 33–52.
Jesswein, K.R., 2009. An examination of the "texas ratio" as a bank failure model. Academy of Banking Studies Journal 8 (2), 63–73.
Jin, J., Kanagaretnam, K., Lobo, G., 2011. Ability of accounting and audit quality variables to predict bank failure during the financial crisis. Journal of Banking and Finance 35, 2811–2819.
Jones, F., 1987. Current techniques in bankruptcy prediction. Journal of Accounting Literature 6, 131–164.
Jordan, D., Rice, D., Sanchez, J., Walker, C., Wort, D., 2010. Predicting Bank Failures: Evidence from 2007 to 2010. Working paper.
Kosmidou, K., Zopounidis, C., 2008. Predicting US commercial bank failures via a multicriteria approach. International Journal of Risk Assessment and Management 9 (1/2), 26–43.
Martin, D., 1977. Early warning of bank failure: a logit regression approach. Journal of Banking and Finance 1, 249–276.
Mayes, D., Stremmel, H., 2014. SUERF Studies. The Effectiveness Of Capital Adequacy Measures In Predicting Bank Distress. SUERF – The European Money and Finance Forum.
Mehran, H., Thakor, A.V., 2011. Bank capital and value in the cross-section. Review of Financial Studies 24 (4), 1019–1067.
Ng, J., Roychowdhury, S., 2014. Do loan loss reserves behave like capital? Evidence from recent bank failures. Review of Accounting Studies 19, 1234–1279.
Pettway, R., Sinkey, J., 1980. Establishing on-site banking examination priorities: an early warning system using accounting and market information. Journal of Finance 35 (1), 137–150.
Sinkey, J., 1975. A multivariate statistical analysis of the characteristics of problem banks. Journal of Finance 30 (1), 21–36.
Thomson, J., 1991. Predicting bank failures in the 1980s. Economic Review of the Federal Reserve Bank of Cleveland 27 (1), 9–20.
Torna, G., 2010. Understanding Commercial Bank Failures in the Modern Banking era. Working paper.
Wheelock, D., Wilson, P., 2000. Why do banks disappear? The determinants of U.S. bank failures and acquisitions. Review of Economics and Statistics 82 (1), 127–138.