Applied Economics and Finance
Vol. 2, No. 2; May 2015
ISSN 2332-7294 E-ISSN 2332-7308
Published by Redfame Publishing
URL: http://aef.redfame.com

Modeling Banks' Probability of Default

Xiaoming Tong[1]

[1] Global Credit Market, Citi Group, New York, USA

Correspondence: Xiaoming Tong, Global Credit Market, Citi Group, New York, USA

Received: February 3, 2015 Accepted: February 26, 2015 Available online: March 23, 2015

doi:10.11114/aef.v2i2.739 URL: http://dx.doi.org/10.11114/aef.v2i2.739

Abstract

The unprecedented financial crisis of 2008-2009 has called attention to limitations of existing methods for estimating the default risk of financial intuitions. Over the past decade, we have had considerable success at predicting default and credit relative value using Merton-type structural models and Hybrid Probability of Default models. However, generating accurate model-based estimates of default probabilities (PDs) for financial firms has proven difficult. To address this need, I built and tested a time-adaptive statistical model that predicts the default probabilities of banks. The model is a logistic regression whose input variables are selected based on their past effectiveness at predicting bank failures and whose inclusion in the model and weights are to be updated quarterly. Model performance at discriminating between defaults and non-defaults was evaluated for horizons of one to five years using a sequence of annual walk-forward out-of-sample tests from 1992 to 2012. I tested the ability of the model to predict absolute default rates out to five years and, except for underestimating the high bank default rates during the credit crisis, the models perform well at estimating the annual bank default rates. Because most default models provide little benefit over agency ratings for low-rated credits, I examined the performance of the model to Kroll agency ratings only for those banks rated above single-B-minus or above single-C-minus. Although default predictions from agency ratings fall off rapidly for banks rated at or above single-B and single-C, the time-adaptive statistical model predictions deteriorate far less. Accuracy at predicting bank defaults using agency ratings decreases to near chance at a prediction horizon of five years, but the time-adaptive statistical model continues to perform well above chance at all horizons. I also present a detailed analysis of the contributions of financial variables to model outputs by year (2000-2012) and tenor (1-5 years) and evaluate the consistency of variable contributions over time. The model performs favorably at predicting defaults, even relative to the best non-financial corporate default models, with a 97% accuracy ratio (AR) at one year prior to default, and decreasing, but still above-chance predictive power out to five years. I find that banks' quality of assets and return on equity are most important for predicting near term defaults, giving way at longer horizons to operating income and the yield on earning assets.

Keywords: bank default, credit risk, default risk

1. Introduction: The Bank Default Model

Over the past decade, we have had considerable success at predicting default and credit relative value using Merton-type structural models, such as Moody's/KMV model (Vasicek, 1988; Kealhofer, 1999) and Citi's Hybrid Probability of Default (HPD) model (Sobehart and Keenan, 2002; 2003). However, generating accurate model-based estimates of default probabilities (PDs) for financial firms has proven difficult. Some reasons for this are financials' high levels of leverage, the relative opacity of their assets and liabilities, potential support from governments, extreme risk of "tail events" and regulatory changes. The numerous bank failures amid the financial crisis of 2008-2009 and the subsequent ratings downgrades of many financial firms have highlighted limitations of agency credit ratings and current credit models to anticipate defaults for financial firms. During the crisis, many banks went from apparent solvency to default in a very short period of time presumably reflecting the particular sensitivity of financial institutions and insurance companies to sudden declines in investor confidence. Although the credit ratings of financial firms are concentrated in the investment grade range, results from Vazza and Kraemer (2012) in Figure 1 demonstrate that despite their higher credit ratings, financial firms have a faster and steeper path to default than their non-financial counterparts. Investors are becoming increasingly interested in better assessing and managing their credit exposure to financial institutions. Also, the U.S. Office of the Comptroller of the Currency (OCC), in accordance with the Dodd-Frank Act, has published final rules (Department of the Treasury, 2012) that remove references to credit ratings from its regulations pertaining to investment securities, securities offerings, and foreign bank capital equivalency deposits.[1] Amid this backdrop, the development of accurate models for assessing bank credit risk appears critical both for managing exposure to financial firms and for compliance with Federal regulations.

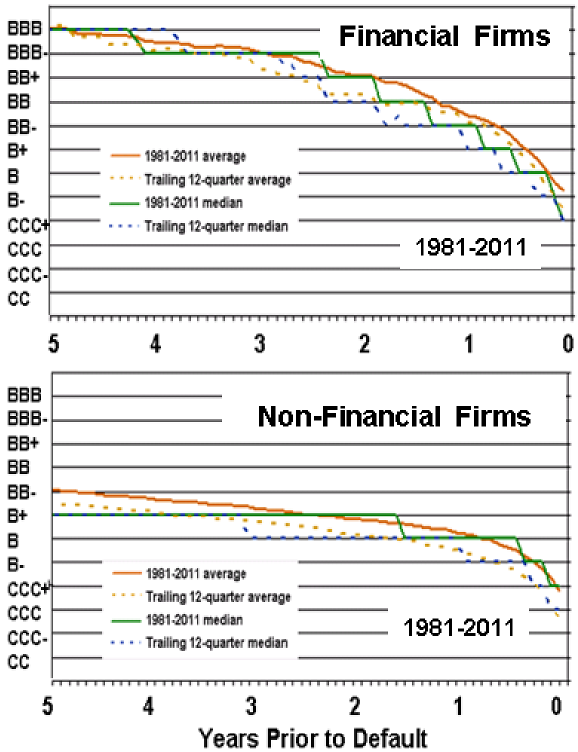


Figure 1. Paths into Default for Financial and Non-Financial Firms from Vazza and Kraemer (2012)

Source: Standard & Poor's

For the development of the time-adaptive statistical model that predicts PDs of banks, I used information contained in banks' financial statements as published by the U.S. Federal Deposit Insurance Corporation (FDIC). As of March 2013, there were 7,019 depository institutions in the US reporting to the FDIC with total liabilities of $12.8 trillion.[2] Potential inputs to the models are financial ratios found, in preliminary analyses, to be effective in forecasting future bank failures. A series of models predicting default at one- to five-year horizons are computed annually, and the outputs are predictions of annual marginal default probabilities for each bank from one to 30 years. I back-tested the model's ability to predict defaults of US depository institutions between 2000 and 2012 using bank data since 1992. For those studies, I evaluated model performance using a walk-forward procedure. That is, to estimate default risk in any test year, I use only information before that year to select model variables and calibrate the model coefficients.

The bank default model is represented in Figure 2. The left panel shows the functional form of the model, a logistic regression, and the most recent coefficients for the one-year model, last updated in 2013 using data up to the end of 2012. As described further below, the variables for each model are chosen based on their relative Bayesian Information Coefficients (BIC), the measure of information contribution, and criterion for inclusion in the model.[3] The middle panel shows how the individual variables are fit to linear regressions and summed prior to input into the non-linear function, $\O(z)$. I built models to predict defaults over yearly horizons from one to five years, each assuming survival (i.e.,

---

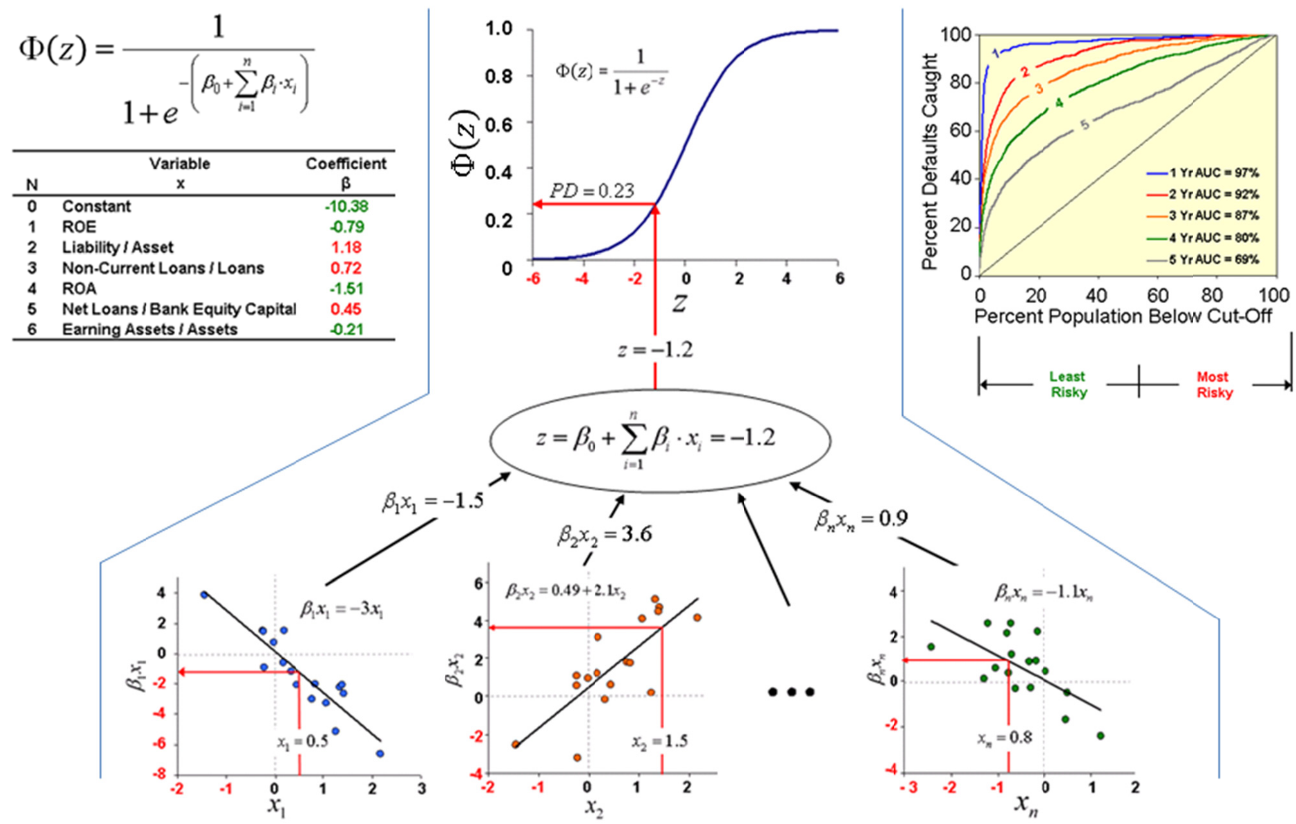
[1] Section 939A of the Dodd–Frank Act requires federal agencies to review regulations that require the use of an assessment of creditworthiness of a security or money market instrument and any references to, or requirements in, those regulations regarding credit ratings. Section 939A then requires the agencies to modify the regulations identified during the review to substitute any references to, or requirements of, reliance on credit ratings with such standards of creditworthiness that each agency determines to be appropriate.

[2] Information about aggregate bank sector size obtained from the FDIC "Statistics on Banking", which is accessible online at http://www2.fdic.gov/sdi/sob/.

[3] That is, as described below, I used the BIC as a criterion for variable selection, with those variables with the highest individual BIC chosen first.

non-default) up to the year of prediction. The panel at the right in Figure 2 shows Cumulative Accuracy Profile (CAP) curves for the five annual prediction horizons. The CAP curves measure the extent to which the model scores serve to separate defaulters (tendency to have high values of $Ø(z)$) from non-defaulters (likelihood of lower values of $Ø(z)$).

Figure 2. Logistic Regression Model for Bank Defaults. Left: Functional form of Logistic Regression and Most Recent Set of Input Variables; Middle: Illustration of How Individual Variables Feed the Non-Linear Regression; Right: CAP Curves for Models to Predict 1- to 5-Year Defaults



The model performs well at identifying the riskiest banks. For example, the right panel of Figure 2 demonstrates that the 10% of banks with the largest values of $Ø(z)$ include 94% of the banks that defaulted within one year.[4] Although performance drops monotonically when predicting defaults for each subsequent year from years two to five (assuming survival to the start of each year), the models perform significantly above chance for all years. That is, for models predicting default in years two, three, four, and five, the10% of banks with the largest values of $Ø(z)$ include 80%, 68%, 55%, and 40% of the defaulting banks in the sample.

In this report, I describe the construction of the model including the method of variable selection, walk-forward validation, and a detailed discussion of model performance. I also describe further features of the model. In particular, I show how values of $Ø(z)$, output from the bank model, are mapped to historical default rates to estimate physical default probabilities. I also describe further validation studies and compare model performance with estimations of bank risk derived from agency credit ratings. Finally, I present detailed studies of the contributions of variables, highlighting which variables are most highly predictive of bank defaults and the periods over which they are most effective.

2. Constructing the Bank PD Model

2.1 Financial Variables as Predictors of Default

Since the pioneering work of Beaver (1966) and Altman (1968), financial modelers have realized that certain financial ratios are highly predictive of a firm's future default. The same is true for banks. For instance, I found that banks with low, especially negative, return on equity (ROE) are much more likely to default. Intuitively, banks with low or negative profitability will likely struggle to pay their liabilities on time and will have difficulty finding additional funding. To

[4] If the model performed at chance, only 10% of the defaulters would be included in the 10% of the sample with the highest values of Ø(z).

illustrate this effect, each panel in Figure 3 displays normalized distributions of ROEs for defaulting and non-defaulting banks. Distributions are shown for one-, two-, three-, and four-year horizons in successive panels.[5] Inspection of Figure 3 reveals that banks with low ROE are much more likely to default than those with high ROEs. Also, the predictive power of the ROE as regards default decreases with increases in the time horizon. That is, the distributions of ROEs from defaulting and non-defaulting banks are clearly apart from each other at one- and two-year horizons, but those differences narrow, becoming very small at four years out. I ran t-tests on the differences between the distributions of ROEs for defaulting and non-defaulting banks cease to be significant over four years out.

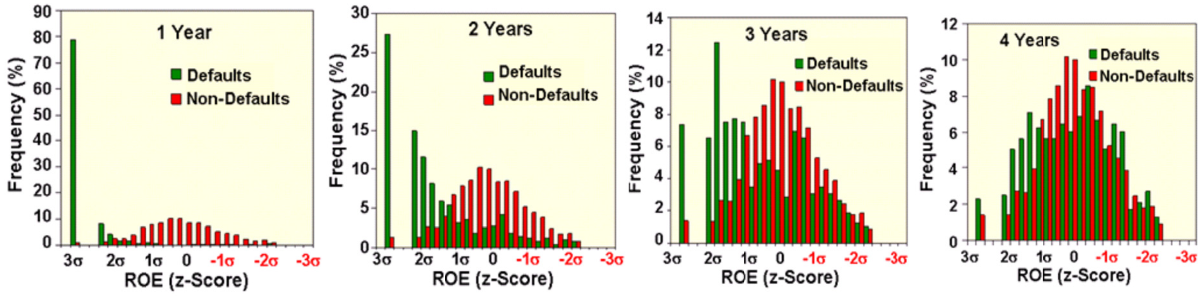


Figure 3. Distributions of Normalized ROEs for Defaulting and Non-Defaulting Banks at One-, Two-, Three-, and Four-Year Horizons

Source: FDIC

A similar testing procedure as illustrated in Figure 3 for ROE revealed other financial ratios that are useful for default prediction. These include firms' leverage ratios, ratios of non-performing to performing loans, and net loans to bank capital, to name a few. A challenge in predicting default is to select an appropriate set of variables and combine them appropriately in a multivariate model. To do this, I employed a walk-forward logistic regression technique. The logistic regression function (described in the following section) is commonly used for predicting variables with binary outcomes, particularly when the inputs are non-linearly related to the desired output. The walk forward method constructs a new model each year from the candidate variable set, while adding the data from the previous year to the development sample. For variable selection in each new model, I use an automated procedure called *forward stepwise selection,* which is explained in detail below.

2.2 Logistic Regression

For those unfamiliar with logistic regression, I describe that method briefly in this section. Logistic regression has similarities to the more familiar multiple linear regression method, but involves an extra step, the logistic transform. I illustrate this graphically for a set of hypothetical input variables in Figure 4. The application begins with selection of a set of candidate financial variables, denoted x_i, i=1,□…, n. The inputs, x_i, could be financial ratios or other quantities. The lower portion of Figure 3 depicts how values of hypothetical input variables (the circles in each plot) are fit by functions, of the form

$$f(x_i) = \alpha_i + \beta_i x_i \tag{1}$$

to derive constants, α_i, and coefficients, β_i, for each input variable.

Then, for a given set of inputs, each x_i is put through its linear transform in Equation 1. For variable x_1 for the example in Figure 3, the constant α_1=0 and the coefficient β_1=-3. Thus, if α_1=0.5 as shown in the figure, $f(x_1)$=-1.5. Hypothetical functions and outputs for x_2 and x_n are also shown in Figure 4.

The resulting outputs of the first stage of the logistic regression, the values $f(x_i)$, are summed at an intermediate stage whose output z can be represented as

$$z = \beta_0 + \sum_{i=1}^{n} \beta_i x_i \tag{2}$$

where

[5] Because financial ratios such as ROEs can have very dispersed distributions, I converted firms' ROEs into standard normal distributions before plotting. This transformation does not change the ordering of firms on the ROE axis.

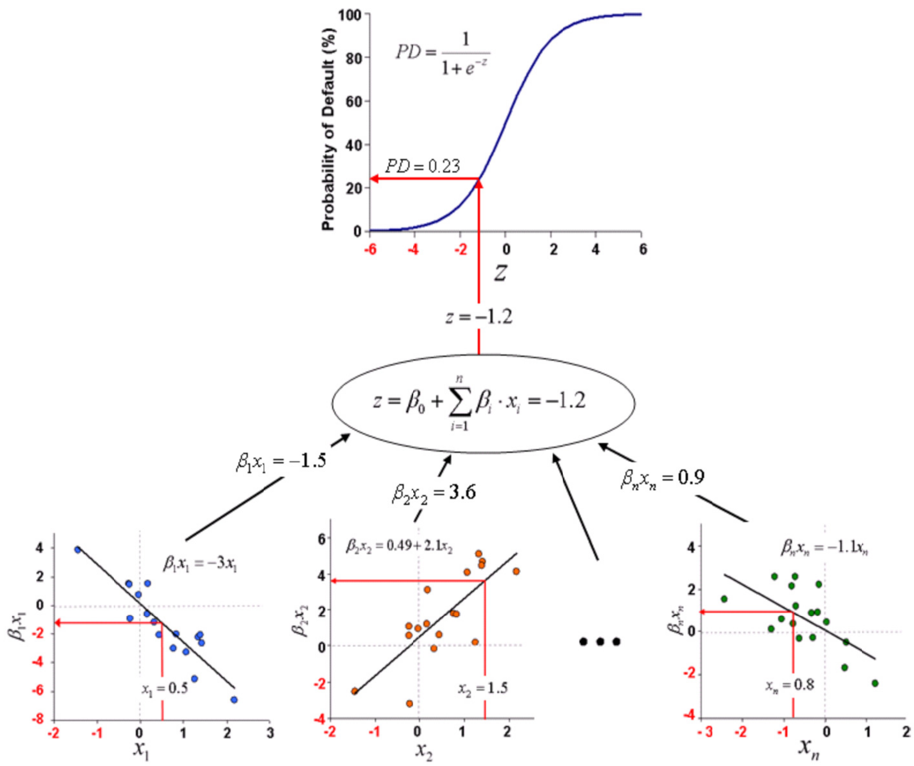
$$\beta_0 = \sum_{i=1}^{n} \alpha_i \tag{3}$$

For the example in Figure 3, the resulting value of z is assumed to be -1.2.[6] The value of z from Equation 2 is then put through the logistic transform that serves to constrain the output of the regression to a value between 0 and 1. For example, for the default model, the resulting PD is given as:

$$PD = \frac{1}{1+e^{-z}} \tag{4}$$

where the resulting value of PD for z=-1.2 is 0.26 or a 26% probability of default over the time frame in question.

Figure 4. Logistic Regression Function. Linear Transformations of Financial Variables (Lower Plots) are Summed at an Intermediate Stage and Put Through the Logistic Transform (Top Graph) Which Converts the Output to a Value between 0 and 1 (0% and 100%)



2.3 Automated Selection of Input Variables

Note that the overall plan is to derive a new model each year, incorporating into the learning sample the data from each successive year's defaulting and non-defaulting firms. Because the factors that influence defaults and their relative contributions may change over time, I chose to use an adaptive procedure for selecting variables for each annual model. I first assembled a set of 20 candidate financial ratios that have been shown to be predictive of subsequent default. Because the distributions of different financial ratios can vary widely, I chose to standardize all input variables via transformation into standard normal distributions before testing their usefulness as inputs to each annual model.

The process of model construction begins with only the logistic function and no variables chosen for inclusion. Then, for each candidate input variable, I build a logistic default model by selecting values of α_i and β_i for each variable that enables the best prediction of default on the development sample. That is for each input variable x_i, I solve for α_i and β_i in the following equation for PD:

$$PD = \frac{1}{1+e^{-(\alpha_i + \beta_i x_i)}} \tag{5}$$

[6] Note that this value of z cannot be deduced from the values shown in Figure 4 as it is assumed to have contributions from variables x3 to xn-1that are not given in the figure.

The variable with the greatest predictive power with respect to default is chosen as the first input variable. As described in further detail below, I chose the Bayesian Information Criterion (BIC) developed by Schwartz (1978) as the measure of predictive power. The BIC measures how well the model fits the data, but also imposes a penalty for having too many variables, thereby guarding against overfitting the data. After selection of the first variable, the process repeated to select a second variable, and so on, until model performance ceases to improve. Once all the variables for the model are selected, the value of the constant β_0 and coefficients $\beta_i (i = 1,...,n)$ for each of the variables are refit to minimize the error in the logistic regression equation:

$$PD = \frac{1}{1 + e^{-(\beta_0 + \sum_{i=1}^{n} \beta_i \cdot x_i)}} \tag{6}$$

$$PD = \frac{1}{1 + e^{-\left(\beta_0 + \sum_{i=1}^{n} \beta_i \cdot x\right)}}$$

N	Variable x	Coefficient β	BIC
0	Constant	-10.38	8,174
1	ROE	-0.79	4,274
2	Liability / Asset	1.18	3,738
3	Non-Current Loans / Loans	0.72	3,522
4	ROA	-1.51	3,465
5	Net Loans / Bank Equity Capital	0.45	3,441
6	Earning Assets / Assets	-0.21	3,437

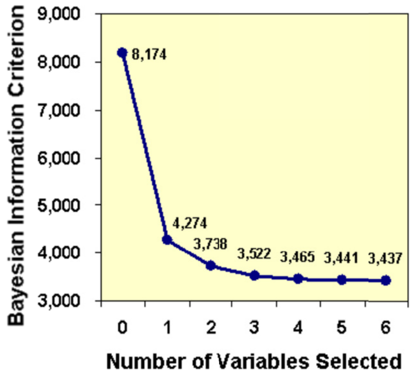


Figure 5. Left: One-Year Bank PD Model Equation, Variables, and Corresponding Coefficients (Green: List Risky; Red: More Risky), With Variables Listed in the Order Selected. Right: Bayesian In formation Criterion (BIC) for Each Successive Variable Selected.

An illustration of the results of variable selection is presented in Figure 5. The top portion of the left panel displays the logistic regression equation, with the table below it listing the input variables to the model in the order in which they are selected. That is, variables are listed in descending order of their predictive power. The BIC values resulting from inclusion of each variable are also displayed. The right portion of Figure 5 is a plot of the BIC values that result from the inclusion of each variable. For instance, the model starts with only a constant term whose BIC value is 7,459. The variable selection procedure determined that banks' return on equity (ROE) provides the largest predictive power of all candidate variables, and its inclusion in the model achieves a BIC of 4,092. After selection of the ROE, the procedure is run again, picking the Liability/Asset ratio as the best of the remaining candidate variables, bringing the BIC down to 3,568. This procedure continued until the BIC could no longer be decreased. At that point, six variables had been selected and their corresponding coefficients appear in the left table of Figure 5.

2.4 The Term Structure of Bank PDs

The method I have described can be used to predict defaults over one- to five-year horizons. However, some applications (e.g., long-term investment portfolios) require estimation of the term structure of PDs over longer periods. My approach to extending the term structure of bank PDs for terms beyond five years is to use long-term annual average marginal default rates determined from historical data on bank defaults.

Construction of PD term structures begins by using the set of five logistic regression models, each developed for the marginal default rate between successive years over a period from one to five years. That is, let PD_t denotes the model designed to predict bank defaults *t* years from now, conditional on the given banks surviving to year *t*-1. That is, for years *t* =1,..., 5, PD_t is the conditional logistic regression model where

$$PD_t = \frac{1}{1 + e^{-\left(\beta_{t,0} + \sum_{i=1}^{n} \beta_{t,i} \cdot x_{t,i}\right)}} \tag{7}$$

Then, for each bank *j*, the probability of default in year *t* assuming survival to year *t-1* is given by

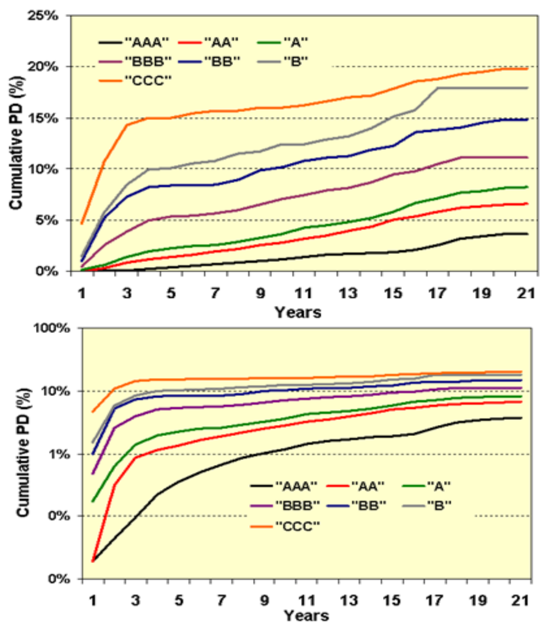
$$PD_{t,j} = \frac{1}{1+e^{-\left(\beta_{t,0}+\sum_{i=1}^{n}\beta_{t,i}\cdot x_{t,i,j}\right)}} \tag{8}$$

Note that because I fit a separate model for each year, the variables selected and the coefficients $\beta_{j,t}$ will, in general, be different for each year. Let, $CPD_{t,j}$ be the cumulative probability of default for bank j from time t=0 to t years. Then, the cumulative probabilities for bank j over horizons from t=1 to T years can be determined from their annual PD$_s$ as:

$$\begin{aligned} CPD_{1,j} &= P_{1,j} \\ CPD_{2,j} &= CPD_{1,j} + (1-CPD_{i,1})\cdot P_{2,j} \\ &.... \\ CPD_{T,j} &= CPD_{T-1,j} + (1-CPD_{T-1,j})\cdot P_{T,j} \end{aligned} \tag{9}$$

The procedure for calculating marginal PD$_s$ beyond five years is illustrated in Figure 6. First, I construct a map between one-year PD$_s$ and Standard & Poor's rating categories. This is made possible using a map that I derived between average probabilities of default for commercial and industrial firms from HPD model (Sobehart and Keenan, 2003) and their corresponding agency ratings.[7] For example, the left panel of Figure 6 illustrates a mapping between one-year PDs from the HPD model to rating categories calibrated using data of all U.S. banks between 1982 and 2012. Using this map, I can assign an implied rating to each bank that corresponds to its current one-year PD from the logistic regression model. Then, for a given bank, I combine its term structure of cumulative default rates from one to five years with the marginal annual default rates reported by Moody's from its imputed credit rating from six to thirty years. That is, I assume each bank's conditional PD beyond five years follows the long-term historical values for its implied rating category. A resulting set of stylized bank annual cumulative default rates by implied whole letter rating categories appear in the right panels of Figure 6. The top panel shows cumulative default rates on a linear PD scale, whereas the lower plot shows those same data in logarithmic PD units. Notice that, as expected, average cumulative default rates for any given tenor increase with decreasing rating categories.

Bank Model 1-Year PD		Implied Rating
Min	Max	
0.00%	0.03%	AAA
0.03%	0.05%	AA+
0.05%	0.06%	AA
0.06%	0.08%	AA-
0.08%	0.10%	A+
0.10%	0.16%	A
0.16%	0.23%	A-
0.23%	0.37%	BBB+
0.37%	0.61%	BBB
0.61%	0.84%	BBB-
0.84%	1.10%	BB+
1.1%	1.3%	BB
1.3%	1.6%	BB-
1.6%	2.1%	B+
2.1%	2.8%	B
2.8%	3.4%	B-
3.4%	4.0%	CCC+
4.0%	6.3%	CCC
6.3%	14.0%	CCC-
14%	31%	CC
31%	67%	C
67%	100%	D



7 That rating map is constructed using PDs from the HPD model for non-bank corporate firms. Then firms are ranked with respect to their model PDs and assigned to rating categories that replicate the number of firms in each rating category in the sample. Finally, implied ratings for U.S. banks are assigned based on their inclusion within PD boundaries determined for each rating category.

Figure 6. Left: Mapping From One-Year PD to Imputed Risk Category; Right: Term Structure of Average Cumulative Bank PD For Each Implied-Rating Category on Linear (Top) and Logarithmic (Bottom) Scales

3. Walk-Forward Backtesting

I back-tested the model by constructing an annual series of models of the bank models using all available US bank data from 1992 to 2012. The number of non-defaulting banks and defaulting banks in the sample by year is given by the green bars (left axis) and red bars (right axis) in Figure 7. Notice that there were roughly 14,000 banks in the sample in 1993, but that number declined to around 7,000 by 2012. Also, there are three apparent waves of defaults: one in the early 90s, a small one around the year 2000, and a surge of bank failures during the recent financial crisis.

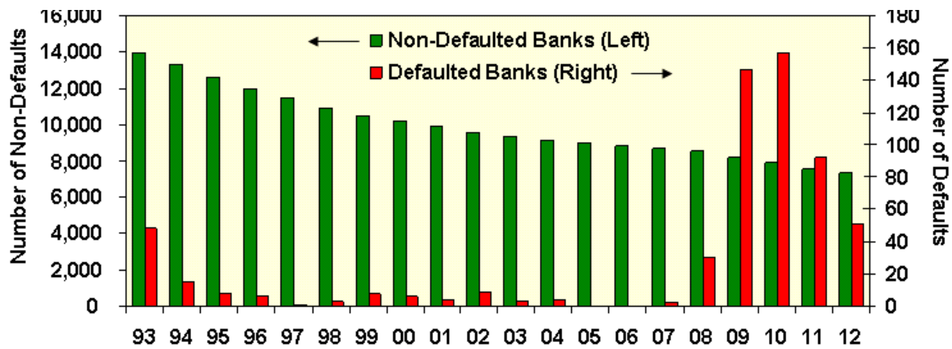
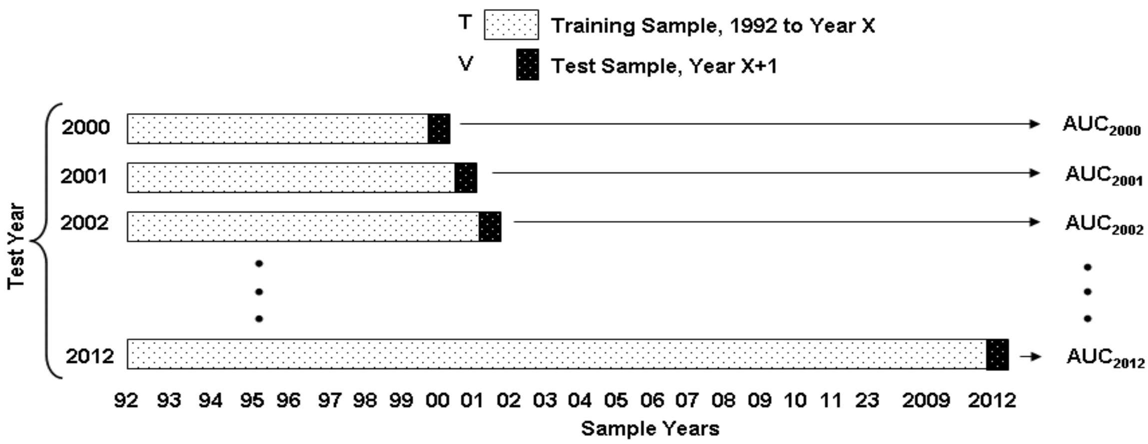


Figure 7. Number of Non-Defaulting Banks (Green Bars, Left Axis) and Defaulting Banks (Red Bards, Right Axis) Banks by Year in the Development Dataset

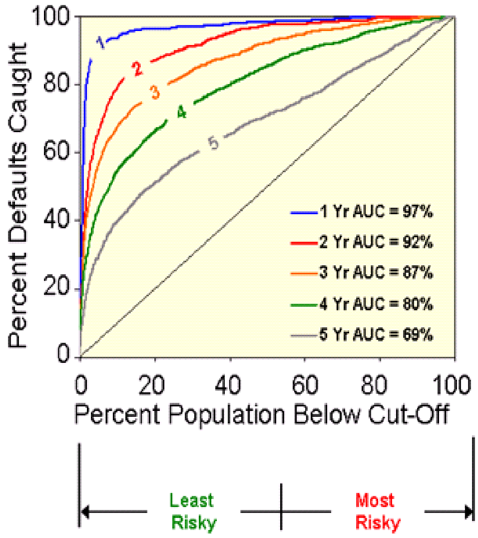
In order to determine out-of-sample performance of the model, I used a walk-forward procedure as illustrated in Figure 8 for the one-year model. The test set is sufficiently large, with a total of 499 defaulters out of 11,114 distinct banks, to provide a strong test of model performance. Because the model needs a minimum number of years of data for development, data from the years 1992 through 1999 were used to construct the first annual model (select variables and calibrate the weights) for each horizon for one to five years. The one-year model for 1999 was then used to score all non-defaulting banks at the beginning of 2000 and its ability to predict defaults in 2000 was determined. Models for year two through five used only banks that had survived to the model year to score for prediction. Thus for the two-year model, firms surviving until 2001 were scored with its 1999 model, and so forth for the longer horizons. To generate the set of models for year 2000 (i.e., used to predict defaults in 2001 to 2005 for one- to five-year models), I added the data from year 2000 to the set from 1992 to 1999. Variables were selected and coefficients determined and the model was tested on the corresponding test sample for the given horizon. That procedure was repeated annually until 2012. Of course, from models at horizons longer than one year, testing could only be done to year 2012 minus the horizon year. I adopted the walk forward procedure because it most realistically estimates the performance of the model as it will be deployed in practice.

Figure 8. Illustration of the Walk-Forward Development and Testing Procedure for the One-Year Models: A New Model is Developed Each Year from 1999 to 2011 Using Data From All Previous Years and Tested on Defaulted and Non-Defaulted Bank in Each Subsequent Year from 2000 to 2012. For Models with Two- to Five-Year Horizons, Test Samples Consisted of Firms Surviving Until Year X+2 to X+5, Respectively.



To evaluate model performance at separating banks that will default from non-defaulters, I generated Cumulative

Accuracy Profile (CAP) Curves for the one- to five-year model horizons. The cumulative resulting CAP curves for test years 1999-2012 are displayed in Figure 9. For example, to generate the one year-curve (blue line), I first rank all banks over the entire 13-year test period from highest to lowest by their one-year PDs from the models. Then, for successive intervals in the ranked population I calculate the cumulative fraction of defaulting banks contained within that interval. The interpretation of CAP curves is straightforward; for any criterion, the fraction of defaulters caught above the population percentile is measures the discriminatory power of the model. For example, the CAP curve for the one-year model at the 10% population criterion caught 94% of the banks that defaulted within the following year over the period from 1999-2012. The higher and steeper the CAP curve over the diagonal chance line, the better the model is at discriminating defaulters from non-defaulters. The table at the right in Figure 9 displays values of the CAP curves for each of the model horizons for various values of the population cut-off. The left-most values in the table show that the 10% of banks ranked riskiest by the one- to five-year models capture 94%, 80%, 68%, 55%, and 40% of the defaulting banks, respectively. Not surprisingly, those data reveal that the power of the models decline as the horizon extends beyond one year, but even the five-year model is performing well above chance, capturing 40% of the banks that default in the fifth year after model development and scoring. Finally, it is important to note that even though the models are only regenerated on an annual basis, the financial data from the banks is available to update bank default scores on a quarterly basis and that is how the model will be used in practice.



Percent Defaults Caught				
	Population Inside Cut-Off			
	10%	20%	30%	50%
One Year	94	96	97	98
Two Year	80	87	91	97
Three Year	68	78	85	92
Four Year	55	67	74	86
Five Year	40	52	60	71
Chance	10	20	30	50

Figure 9. Left: CAP Curves for Predictions of Bank Defaults for One- to Five-Year Models using Walk-Forward Testing from 1999 Through 2012; Right: Values of the One- to Five-Year CAP Curves at Critical Thresholds, with Corresponding Values from the Chance Line Also Shown

From a risk management perspective, the most relevant horizon for prediction is at one year. Thus, if a bank survives for that one year, the next year's model can be used to assess its subsequent risk. Still, there are applications for which multi-year estimates of losses and portfolio relative value are of interest. These include buy-and-hold portfolios of bank obligations, such as structured products. For example, if one holds a portfolio of bank TRUPS (trust preferred securities) with five years of remaining maturity, they may wish to estimate-five year portfolio losses. For this type of application, it is important that the absolute PD levels be accurate. The CAP curves, because they rank PDs, assess only the relative accuracy of the models.[8] Indeed, the models do specify absolute PD levels and I can assess their accuracy using the reliability plots in Figure 10. To construct the plots in Figure 10 I separated all banks into bins by 5% PD increments, and plot each bin's average predicted PDs on the horizontal axis and the realized rate of defaults on the vertical axis. The interpretation of reliability plots is as follows. For example, the one-year plot includes the point (27% predicted, 31% obtained), which means for all the banks assigned one-year PDs between 25% and 30%, 31% of them actually defaulted within the following year. A perfect model would have all points falling on the diagonal line for which predicted PD and realized default rates match exactly. Error bars at two standard deviations for the realized default rates are also shown in each plot.

[8] For example, if one multiplies all PDs by 10 the CAP curves will not change, but the absolute PD levels implied by the models will be too large.

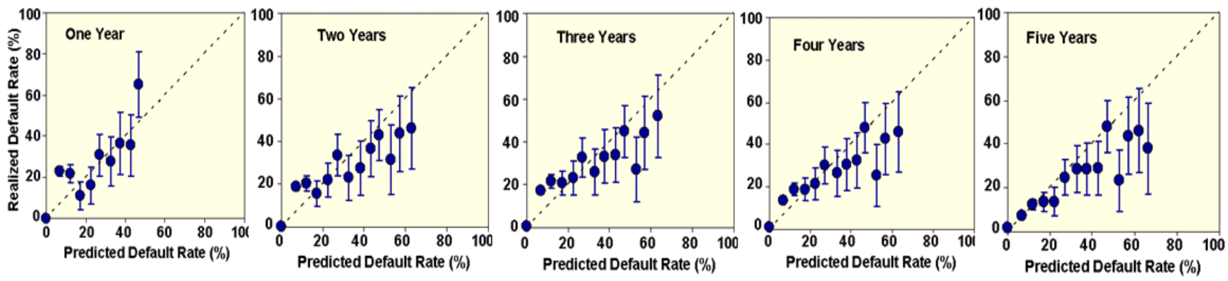


Figure 10. Plots of Predicted and Obtained Bank Annual Default Rates from Models for One to Five Years Out. Error Bars are Two Standard Deviation Bands

The plots in Figure 10 indicate that the default probabilities generated by the model are reasonably accurate at predicting default rates for banks over multi-year horizons. With respect to the two standard deviation bars, most data predictions do not differ significantly from the diagonal "perfect model" line. However, a notable exception is that the bank model typically underestimates the default rates for the second- and third-highest bins (i.e., the high default 60%-70% bins). Further analysis revealed that the model under-predicted the sudden surge of defaults during the financial crisis of 2008 and 2009. Consider the left panel Figure 11 which displays the historical annual high yield corporate default rates (left axis) and U.S. bank default rates (right axis) from 1993 through 2012. Notice that the high yield default rates varied substantially over the period, with high rates early in the century. The banks had been relatively safe before 2008, with an average annual default rate of only 0.06% and even the maximum during that period is only 0.34%. The right panel of Figure 11 plots average predicted and realized annual default rates from the one-year bank model. The bank default models that are constructed annually did not predict well the overall bank default rate in 2008 and 2009, the years of high bank defaults. More generally, the plot reveals that PD levels from the bank model tend to trail observed annual PD rates by one year. Note that the financial data for U.S. banks are published quarterly by the U.S. Federal Deposit Insurance Corporation (FDIC). Thus, in practice, I plan to update the model quarterly, potentially minimizing the lag in accurately predicting annual default rates.

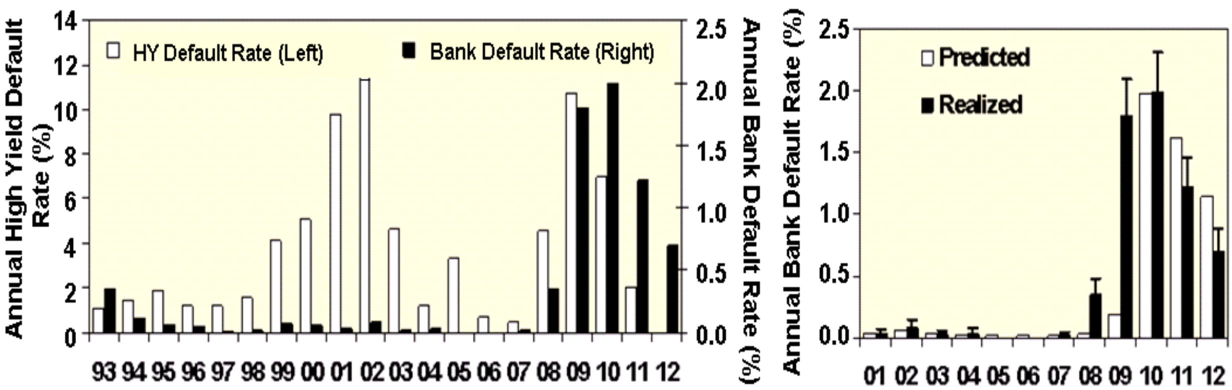


Figure 11. Left: Historical Annual High Yield Default Rate and Bank Default Rates From 1993 to 2012. Right: Model Predicted Annual Bank Default Rate and Realized Bank Default Rates, 2001–2012 with Two Standard Deviation Error Bars

3.1 Converting Model Scores to Default Probabilities

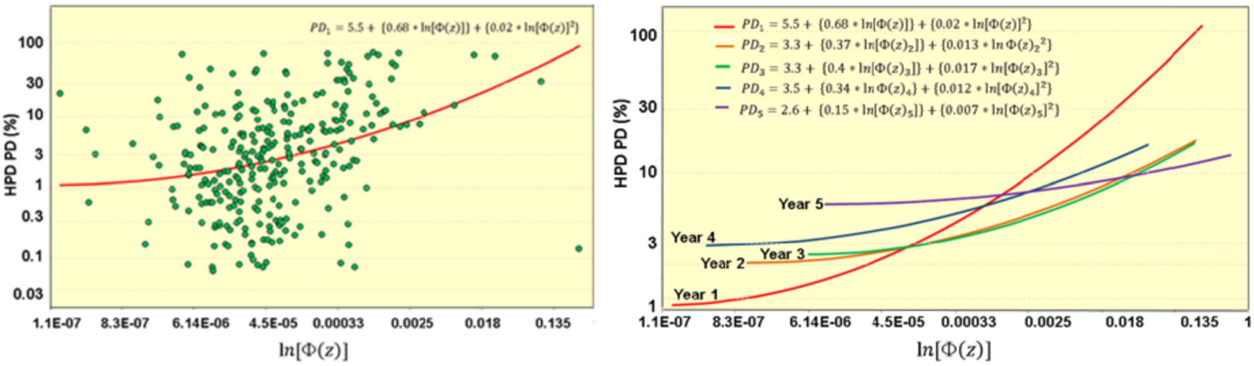
I previously showed that values of *Ø*(*z*) from the model are highly correlated with default probability (see right panel of Figure 12). That is, the model appears to perform well at ranking the relative default risk of U.S. banks. Although I attempted to link outputs of the model (i.e., values of *Ø*(*z*)) to actual physical default probabilities, the resulting values proved less than satisfactory. Accordingly, in this section, I link values of *Ø*(*z*) from the bank model to default probabilities from Hybrid Probability of Default (HPD).

My approach to transforming values of *Ø*(*z*), for $i = 1,...,5,$ where i indicates model for a given default year contingent upon survival to year $i = 1$, is straightforward. For those banks that have PDs from HPD model, I plot HPD PDs versus values of $\ln\phi(z)_i$ from the bank model as shown in the left panel of Figure 12 for the one-year model (i.e., $i = 1$). Then I fit the points with a second- order polynomial of the form

$$PD_i = a_i + \left[b_i \cdot \ln \phi(z)_i \right] + \left[c_i \cdot \ln \phi(z)_i^2 \right] \tag{10}$$

as shown by the red line in the figure for which $i = 1$,and $a_1 = 5.5$, $b_1 = 0.68$, and $c_1 = 0.02$. It is important to note that I impose monotonicity on the function in Equation 10 to ensure that the conversion from $\ln \phi(z)$ to *PD* does not change the ordering of banks as regards their default risk. Thus, the transformation in Equation 10 merely serves to transform model outputs to physical PDs and does not alter the CAP curves shown in Figure 2.

Figure 12. Mappings Between Bank Model Outputs, $\ln \phi_i(z)$ and HPD PDs. Left: Best Fit Order 2 Polynomial of HPD PDs to $\ln \phi_1(z)$; Right: Best Fit Order Polynomial 2 to $\ln \phi_i(z)$ for $i = 1,\ldots,5$ Models



The right panel of Figure 12 shows the mapping from $\ln \phi(z)$ to *PD* function for the one-year model (i.e., $i = 1$) using the red line, along with the functions for years two to five. The coefficients of a_i, b_i and c_i for each of the curves are inset in the graph. Finally, note that these mapping functions are only used for the current set of models. Although the mapping in Figure 12 is for December 5th, 2013, this mapping can be updated daily to reflect changes in banks' PDs from the HPD model.

4. Performance of the Bank Model Relative to Agency Credit Ratings

In this section I performed a series of studies to address the extent to which my models offer any advantages over using agency ratings for assessing bank risk. Investors are interested in the ability to predict defaults for credits not already recognized by the agencies as risky. That is, "How well does the model do at predicting defaults for banks with credit ratings above triple-C, single-B, and so on?"

Of particular interest to investors is the riskiness of smaller and/or lower-rated financial firms, particularly savings and loans and bank holding companies. This is because debt from those financial firms is often placed in TRUPs (Trust Preferred securities). Because returns and payouts from TRUPs are highly dependent on defaults and ratings downgrades, those investors are particularly interested in accurate assessments of default probabilities and signals of deteriorating credit quality. To test this, I examined the relative predictive power of ratings by Kroll, who are best known in this space, and my bank default models for savings and loans and bank holding companies.

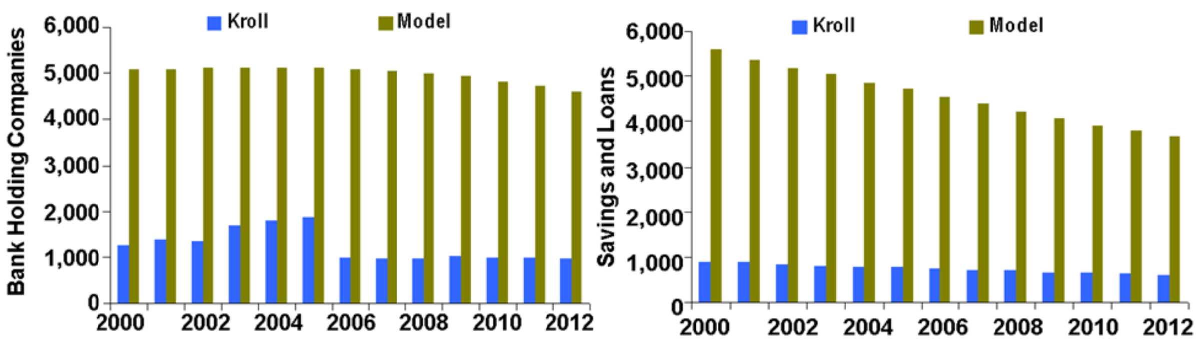


Figure 13. Financial Firms Having Both Kroll Ratings and My Bank Model Scores: Left: Bank Holding Companies; Right: Savings and Loans

Source: Kroll Rating Agency

To test the predictive power of my model versus agency ratings, I first determined those financial firms that have both model scores and Kroll agency ratings. Kroll has three categories of financial firms: bank holding companies (BHC), Savings and Loans (SNL), and Banks (not identified). I obtained Kroll ratings for as many financial institutions as possible over the period from 2000-2012. The number of banks having Kroll ratings and my model scores appear in Figure 13, broken out by bank holding companies and savings and loans. Clearly, it appears that Kroll rates significantly less financial firms than are scored by my bank model. Part of this is because Kroll does not rate new banks within the first three years of their existence. It is also possible that I received only partial data on Kroll bank ratings. Nevertheless, as shown in Figure 13, there are roughly 1,000 firms each year having both Kroll ratings and my model scores, and these are typically the lower rated portion of the financial services firms. Importantly, as shown in Figure 14, there are at least a reasonable number of defaults for testing model, at least when results are aggregated over the 13-year test period.[9]

Figure 4. Number of Defaults for Testing One- to Five-Year Models

	Number of Defaults				
	Horizon (Years)				
	1	2	3	4	5
BHC	61	121	172	215	317
S&L	109	214	312	403	459

Figure 14. Number of Defaults for Testing One- to Five-Year Models

To test the predictive power of my bank PD model versus agency ratings, I first determined those financial firms that have both model scores and agency ratings. The left panel of Figure 15 shows the Kroll rating scale, where ratings range from single-A-plus to default (D). The middle and right panels of Figure 15 display the distributions of bank holding companies and savings and loans by Kroll credit ratings, respectively, for financial firms having both Kroll credit ratings and my bank model scores. Notice that Kroll does not rate many banks or savings and loans at A+ or A-. Of course, there are relatively few very low rated (single-C-plus to single-C-minus) financial institutions as it is very difficult for low-rated financial institutions to survive for long. Notice also that there are fewer savings and loans than bank holding companies in the sample.

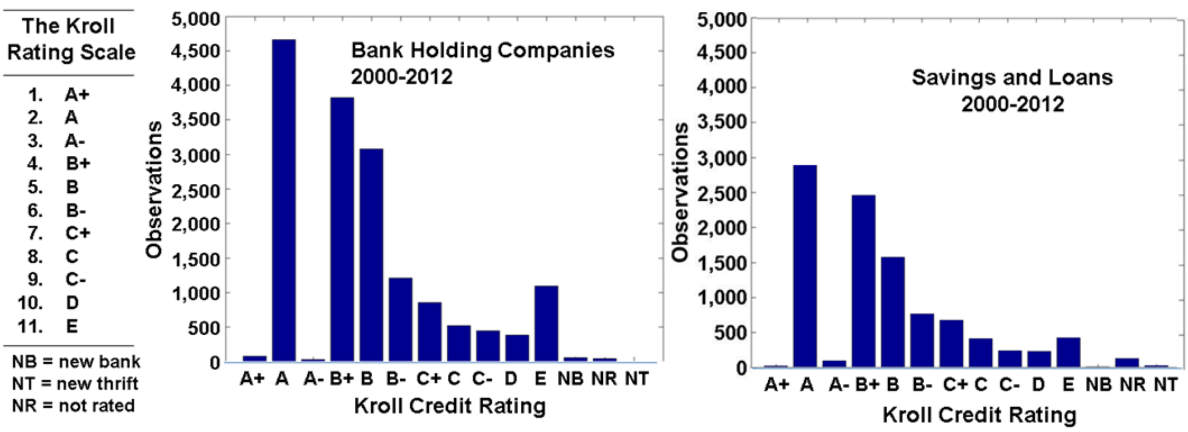


Figure 15. Kroll Rating Scale for Financial Firms (Left) and Distributions of Bank Holding Companies (Middle) and Savings and Loans (Right) with Both My Bank Model Scores and Kroll Credit Ratings

Source: Kroll Rating Agency

Default is necessarily a probabilistic event. That is, one is rarely certain that an obligor will default until its actual occurrence. A useful method for evaluating models' predictive accuracy of probabilistic events is by constructing cumulative accuracy profile (CAP) curves such as that shown in the upper panel of Figure 2. Construction of cumulative accuracy profiles are described in detail in many places (Sobehart and Stein, 2000) and a short description appears in Appendix A. Briefly, to construct a CAP curve for a bank default model, values of estimated risk are first ranked from largest to smallest, with information whether each score is associated with a subsequently defaulted bank

[9] Notice in Figure 14 that numbers of defaults increase with model horizon. This is because I use overlapping windows in counting multi-year defaults.

or a solvent one. Then, I start from the banks with the riskiest scores, say the top 1%. In that top 1% of banks with the largest estimates of default, I calculate the percentage of defaulters in that sample. For example, if a model were assigning scores at random, one would observe only 1% of the defaulters in the top 1% of the population. Any percent of defaulters above 1% would be indicative of the model's predictive power.

I computed CAP curves for predicting bank defaults using Kroll agency ratings and my bank model. Figure 16 displays CAP curves for bank holding companies (top) and savings and loans (bottom) for predicting default in year one to four. The curves for Kroll ratings and my bank model are presented in each graph for comparison. A useful measure of predictive power from CAP analysis is the Area under the Curve (AUC), which is the percentage of the area under each CAP curve. The AUCs for Kroll ratings and my bank model are inset in each plot.

Several features of the data in Figure 16 are of interest. First, it is clear that both my bank default model and the Kroll agency ratings order banks' risk at better than chance levels, even out to four years. It is also evident that predictive accuracy for both models decreases as the year of prediction gets farther out in time.[10] Visual inspection of AUCs in the top and bottom panels is sufficient to conclude that each model's performance for bond holding companies and savings and loans are similar. This is confirmed by values of AUCs listed in tabular form in Figure 17 for each model and year (these same values are inset in each plot in Figure 16). That is, AUCs for BHCs and SNLs within each model vary at most by 4% and often only by 1%.

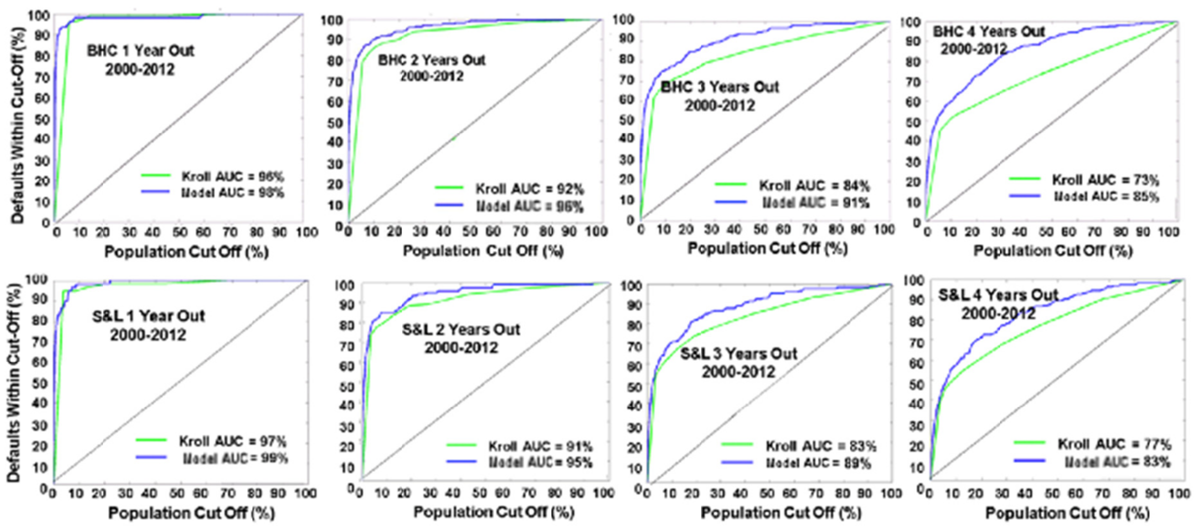


Figure 16. Comparison of Cumulative Accuracy Profile (CAP) Curves for Bank Holding Companies (Top) and Savings and Loans Between Kroll Agency Ratings and My Bank Model from 2000-2012

Source: Kroll Rating Agency

	Horizon (Years)							
	1		2		3		4	
	BHC	SNL	BHC	SNL	BHC	SNL	BHC	SNL
Kroll	96	97	92	91	84	83	73	77
Model	98	99	96	99	91	89	85	83
M-K	2	2	4	8	7	6	8	6

Figure 17. Areas Under the CAP Curves for Kroll Agency Ratings and My Bank Model for Bank Holding Companies and Savings and Loans

Source: Kroll Rating Agency

Figure 16 and Figure 17 also allow comparison of performance between Kroll ratings and my bank model at predicting bank defaults. Again, visual inspection of the CAP curves in Figure 16 is sufficient to conclude that my bank model

[10] Recall that for each model the prediction for each year is dependent on the firm surviving up to the year of prediction. Thus, when predicting defaults for year two, all firms that defaulted in the first year after the date of prediction are excluded from the sample.

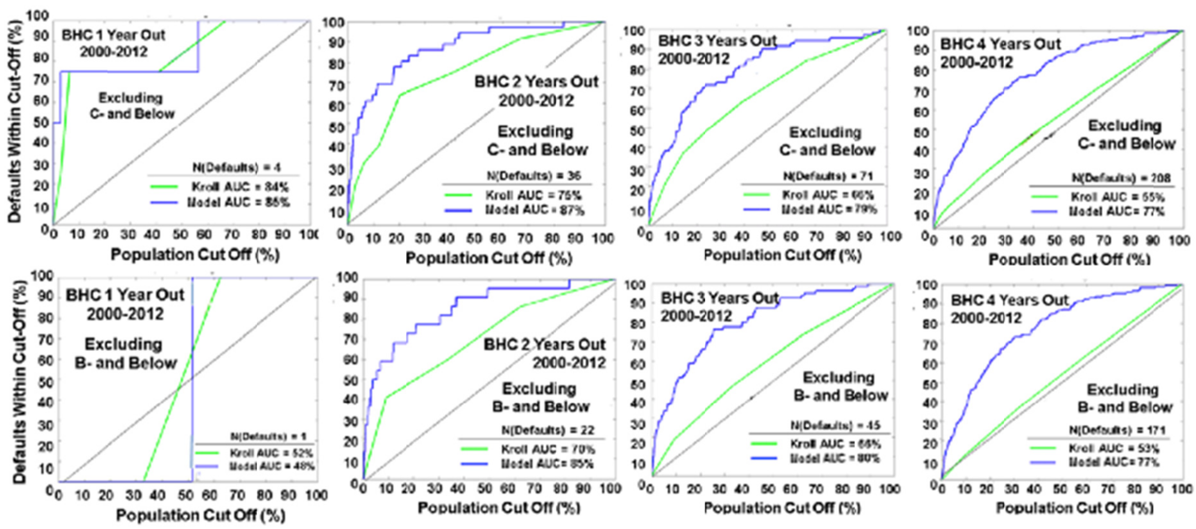
performs better than Kroll ratings, particularly as the time horizon increases. Notably, for one-year predictions, the models are both quite good; AUCs for my bank model are 98% and 99% for BHCs and SNLs, respectively, with AUCs for Kroll ratings are 96% and 97%. Still, as shown in the last row of Figure 17, my model edges out Kroll ratings even at one-year, with that advantage tending to increase with prediction horizon. In fact, for all eight CAP curves in Figure 16. AUCs for my bank default model are greater than those for Kroll ratings.

Despite the success of my bank model in predicting defaults, both investors and traders have remarked the predicting bank defaults at short horizons is not difficult. That is, they claim that bank failures tend to be rapid and fairly obvious. Some suggest that deterioration of banks' credit is reflected in high levels of non-performing loans and loss of investor confidence as evidenced by rapid withdrawal of banks' necessary short term funding. Furthermore, agency ratings are sufficient to capture those aspects of bank performance. That the Kroll ratings and my bank model have extremely high predictive power at the one-year horizon is consistent with that view. However, the predictive power of my bank model at longer horizons relative to agency ratings suggests that my bank model is adding value. I evaluate these issues in the remaining section of this study.

5. Predicting Default for Less Risky Banks

One way to assess the added value of bank default models, given that low-rated financials have already been recognized by agencies and investors as risky, is to exclude the riskiest obligors from the analysis of performance. To that end, in separate analyses, I eliminated all those obligors rated by Kroll's below single single-C (i.e., single-C-minus and below) and all those banks Kroll rates below single-B (single-B-minus and below).[11] I then computed CAP curves and AUCs on those sub-samples as in Figure 16 and Figure 17, respectively. The resulting CAP curves and table of AUCs for Bank Holding Companies appear in Figure 18 and the left table in Figure 19, respectively.[12] Consider first the CAP curves in Figure 18. CAP curves excluding BHCs rated below single-C appear by horizon in the top panels and those excluding BHCs rated below single-B appear in the lower panels. First, notice that only one firm rated below single-C and four BHCs rated below single-B default within one year. This confirms the intuition, stated above, that the predicting bank defaults at short horizons is not difficult given their low agency ratings. That is, the defaults that occur within one-year of risk scoring occur almost exclusively for banks rated below single-C by the Kroll agency.

Figure 18. Cumulative Accuracy Profile (CAP) Curves for Predicting Defaults on Bank Holding Companies for Kroll Ratings and My Bank Default Model. Top: CAP Curves Excluding Credits Rated by Kroll Below Single-C; and Bottom: CAP Curves Excluding Firms Rated Below Single-B, 2002-2012



---

[11] It is important to keep the distinction in mind between the Kroll rating scale in the left portion of Figure 5 and the more familiar scales of Standard and Poor's and Moody's. As a rule of thumb, a Kroll rating of single-C is roughly equivalent to a rating by Standard and Poor's of triple-C.

[12] Because PD score can take any value between 0 and 1, while ratings can only take one of a small number of discrete values, my scores also allow for finer discrimination between institutions than ratings do. This difference also explains why the CAP curves for PD model is a step-function while the CAP curve for ratings is a piecewise linear function. This is particularly evident as the default sample is small as for the one-year horizon model.

	Bank Holding Companies (BHC)							
	Horizon (Years)							
	1		2		3		4	
Exclude Below	C	B	C	B	C	B	C	B
N(Defaults)	4	1	36	22	71	45	20B	171
Kroll	84	52	75	70	66	66	55	53
Model	85	48	87	85	79	80	77	77
Model-Kroll	1	-2	12	15	13	14	12	14

	Savings and Loans (SNL)							
	Horizon (Years)							
	1		2		3		4	
Exclude Below	C	B	C	B	C	B	C	B
N(Defaults)	3	1	27	14	71	45	116	85
Kroll	77	55	77	61	66	59	62	60
Model	82	89	86	80	79	78	77	76
Model-Kroll	5	34	9	21	13	19	15	16

Figure 19. Areas Under the CAP Curves (AUCs) Excluding Firms Rated by Kroll as Single-C-Minus and Below and Single-B-Minus and Below for Bank Holding Companies (Left) and Savings and Loans (Right)

The pattern of CAP curves in Figure 18 is similar for banks rated by Kroll at or below single-C-minus and single-B-minus. First, for all horizons greater than one-year, my bank model outperforms Kroll agency ratings. Also, performance appears to decrease as the time horizon increases from two to four years, but then decreases in performance are greater for Kroll ratings than for my bank model. I note that for predicting defaults at the four-year horizon, the CAP curves indicated that Kroll ratings are nearly at chance, whereas my bank model is performing well above chance. Interestingly, performance of both my bank PD model and Kroll ratings do not appear to change significantly between BHC samples that exclude firms below single-C or single-B.

The features of the CAP curves mentioned above for Bank Holding Companies are presented quantitatively as AUCs in the left portion of Figure 19. That is, except for one-year, already discussed above, AUCs from my bank model are greater than those for Kroll ratings at all tenors, by 8% to 14%. The pattern is similar when excluding either BHCs rated below C or B. That is, performance appears to drop off for both models as prediction horizon increases, but that performance drops off more dramatically for Kroll ratings, being close to chance at the four-year horizon. Also, AUCs for cases where BHCs below B and C are excluded are remarkably similar; model performances do not drop off appreciably when the criterion for inclusion in the sample is raised to B and above from C and above.

The right portion of Figure 19 lists similar measures for Savings and Loans. The pattern of results mimics well those for the BHCs in the left panel:

- My bank PD model outperforms Kroll ratings at all tenors when SNLs rated below C and B are excluded;
- AUCs tend to decrease with increases in year of default prediction, but those decreases, while slight for my bank PD model, are greater for Kroll agency ratings;
- Performance of both my bank PD model and Kroll's ratings are similar when either SNLs below C and B are excluded from the analysis.

The CAP curves for SNLs are remarkably similar to those for BHCs. The CAP curves for Savings and Loans that correspond to those presented in for Bank Holding Companies in Figure 18 are presented for comparison in Figure 20. Note that the pattern of CAP curves for SNLs in Figure 20 is highly similar to those for BHCs in Figure 18.

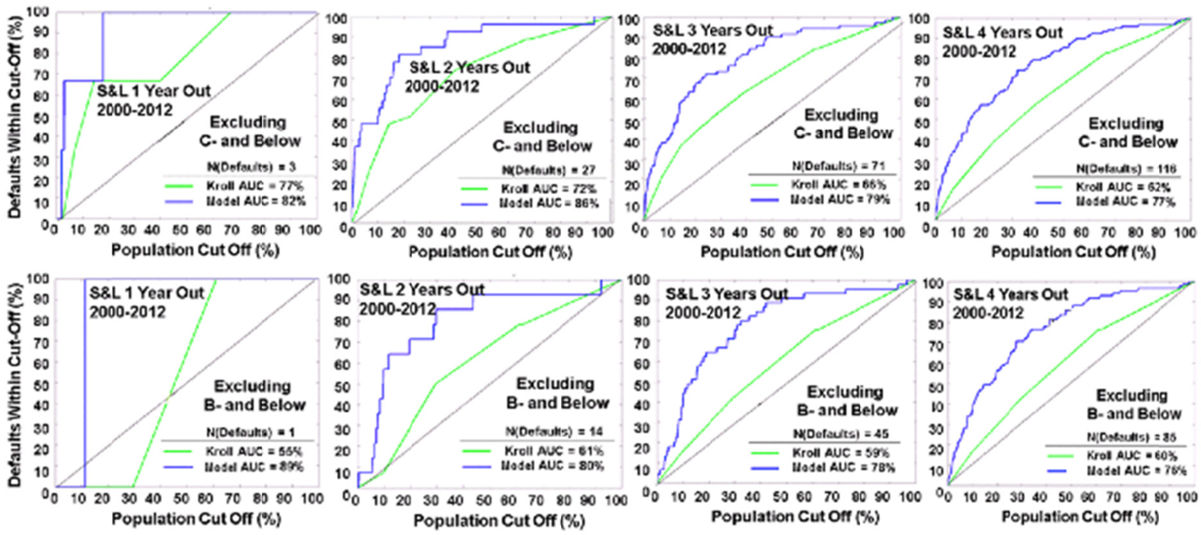


Figure 20. Cumulative Accuracy Profile (CAP) Curves for Predicting Defaults on Savings and Loans for Kroll Ratings and My Bank Default Model. Top: CAP Curves Excluding Credits Rated by Kroll Below Single-C; and Bottom: CAP Curves Excluding Firms Rated Below Single-B, 2002-2012

6. Analysis of Variable Contributions

From the analysis of the CAP curves and AUCs above, it is clear that my bank PD model continues to perform well, even as the year of default prediction moves out. One of the reasons for this is that the bank PD model uses different sets of variables to predict default at each successive horizon. Accordingly, in this section, I detail the contributions of the various candidate variables to the models, both over the time frame of testing from 2000-2012 and for tenors of one to five years.

As described in detail previously, for each test year in the walk-forward model development procedure from 2000-2012, I built models to predict defaults for one-, two-, three-, four- and five-year horizons. To construct each model in each year, I began by selecting first the variable having the largest Bayesian Information Criterion (BIC) when predicting default alone.[13] Then I iterated this procedure, adding variables one-by-one, until no further improvement in overall

[13] The Bayesian Information Criterion was developed by Schwartz (1978).

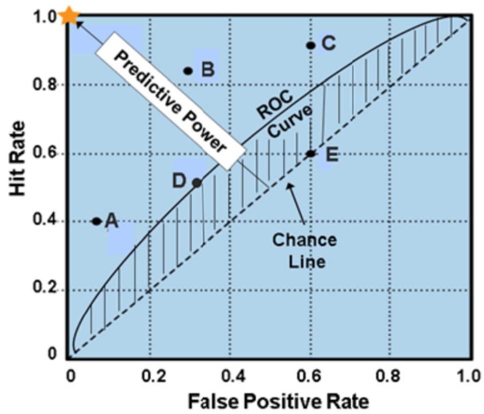
Appendix A: Cumulative Accuracy Profile (CAP) Curves and Receiver Operating Characteristics (ROCs)

The cumulative accuracy profile (CAP) curve and its close relative, the receiver operating characteristic (ROC)

Appendix A: Cumulative Accuracy Profile (CAP) Curves and Receiver Operating Characteristics (ROCs)

The cumulative accuracy profile (CAP) curve and its close relative, the receiver operating characteristic (ROC) are two-dimensional plots where the variable of interest, say the likelihood that the bank defaulted in the following year is plotted against the predictive variable, the ranked magnitude of the bank's risk of default as estimated by a given model.

Figure 29: Top: CAP Curve with Points to Illustrate Various Features; Bottom: Descriptions of Points on the CAP curve



Point	Description
★	Perfect Classification
E	Poor Classification (Equivalent to Guessing)
A	Conservative Classification
B	More Permissive Classification than A, More Conservative than C
C	More Permissive than A and B
D	Worse Classification Than B (Hit Rate is Lower and False Positive Rate is the Same

The CAP curve at the top of Figure 29 illustrates some important features:

- The negative diagonal shown by dashes is the “chance line” where the likelihood of detecting a subsequent default is unrelated to the risk rating from the given model; the “hit” and “false positive” rates are the same;
- The “CAP curve”, the solid curved line in Figure 27, the plot of “hits” versus “false positives” as one goes through the entire set of banks ranked by their predicted credit risk from largest to smallest; and
- The dashed area under the CAP curve, called the “area under the curve” or AUC, is an important measure of performance, whose value is the area under the CAP curve.

The bottom portion of Figure 29 provides descriptions of the points labeled in the CAP curve. For example, the “star” in the plot at (hit, false alarm) = (0.0, 1.0) is indicative of a perfect predictor; all target cases are ranked above the first “false positive” in the sample. That is all defaulted banks larger risk ratings than those that did not default. A point along the chance line, “E,” is also shown. The point “A” is said to plot at a “conservative” criterion, as the false positive rate is very low. Points “B” and “C” illustrate more permissive criteria than “A.” Finally, the point “D” is on the ROC curve for a given level of discrimination. Note that points A, B, and C could intersect the same ROC, which would indicate superior predictability to that shown through point “D.”

Appendix B: Description of Candidate Variables

For clarification, I present more detailed descriptions of some of the candidate variables.

ROA: Net income after taxes and extraordinary items (annualized) as a percent of average total assets.

Non-Current Loans / Loan Loss Allowance: Non-current loans are defined as assets past due 90 days or more, plus assets placed in nonaccrual status. This quantity is divided by loan loss allowance. Loan loss allowance is the amount each bank must maintain in reserve for loan and lease losses to absorb estimated credit losses associated with its loan and lease portfolio (which also includes off-balance- sheet credit instruments).

Earning Assets / Assets: This is the ratio of all loans and other investments that earn interest or dividend income to the sum of all assets owned by the institution including cash, loans, securities, bank premises, and other assets. This total does not include off-balance-sheet accounts.

Net Interest Margin: Total interest income less total interest expense (annualized) as a percent of average earning assets.

Yield on Earning Assets: Total interest income (annualized) as a percent of average earning assets.

Net Operating Income to Assets: Net operating income (annualized) as a percent of average assets.

BIC occurred. A list of the candidate variable set used for model construction appears in Figure 21 and a more detailed description of each candidate variable appears in Appendix B.

Figure 21. Candidate Variables for My Bank Default Models for One- to Five-Year Horizon

Candidate Variables for Bank Models
ROE
Liabilities / Assets
Non-Current Loans / Loans
Net Operation Income / Assets
Earning Assets / Assets
Yield on Earning Asset
ROA
Assets
Non-Current Loans / Loan Loss Allowance
Net Loans / Bank Equity Capital
Annual Default Rate
Assets 90 days Past Due / 30 - 89 Days Past Due
Net Interest Income / Earning Assets
Net Interest Margin

Figure 22. Left: Potential Model Variables and Their Order of Selection for the One-Year Model for 2000 to 2012; Right: Summary of Variable Selection, Displaying Probability of Selection and Average Order if Selected for the One-Year Model

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Variables	Order in Which Selected (1,2,...)												
ROE	1	1	1	1	1	1	1	1	1	1	1	1	1
Liabilities / Assets	2	2	2	2	2	2	2	2	2	2	3	2	2
Non-Current Loans / Loans		5	5	5	5	4	4	4	4	3	2	3	3
Net Operation Income / Assets	4	4	4	4	4	5	5	5	5	6	6	7	
Earning Assets / Assets	3	3	3	3	3	3	3	3	3	4			7
Yield on Earning Asset					6	6	6	6	6	5			
ROA											4	4	4
Assets											5	6	6
Non-Current Loans / Loan Loss Allowance												5	5
Net Loans / Bank Equity Capital													
Annual Default Rate													
Assets 90 days Past Due / 30 - 89 Days Past Due													
Net Interest Income / Earning Assets													
Net Interest Margin													
Num of Vars Selected	4	5	5	5	6	6	6	6	6	6	6	7	7

1 Year Model		
Variables	Probability of Selection	Average Order if Selected
ROE	100%	1.0
Liabilities / Assets	100%	2.1
Non-Current Loans / Loans	92%	3.9
Net Operation Income / Assets	92%	4.9
Earning Assets / Assets	85%	3.5
Yield on Earning Asset	46%	5.8
ROA	23%	4.0
Assets	23%	5.7
Non-Current Loans / Loan Loss Allowance	15%	5.0
Net Loans / Bank Equity Capital	0%	-
Annual Default Rate	0%	-
Assets 90 days Past Due / 30 - 89 Days Past Due	0%	-
Net Interest Income / Earning Assets	0%	-
Net Interest Margin	0%	-

Consider first variables selected for the one-year model. The left panel of Figure 22 lists the variables selected each year from 2000 to 2012 for the one-year model using my walk-forward development and testing procedure. The number associated with each variable is the order in which the variable was chosen based on its BIC. Those variables having no value associated with them in a given year were not selected for that year's model. The table shows that the most important variable for predicting one-year default is firms' return on equity (ROE). The ROE was the first variable chosen for one-year models in every year. The ratio of liabilities to assets is also important for short-term default predictions, being chosen second for all years except 2010. Contributions from other variables are less consistent over time. For example, prior to 2009, earning assets to total assets is the third variable selected for all models, but was not even selected in 2010 or 2011, being selected last in 2012. Conversely, the ratio of non-current loans to loans was not selected at all in 2000, gradually increasing in its importance, such that from 2009 onward, it was either the second or third most important predictor of default. Other variables having contributions to one-year default prediction are the yield on earning assets over the period from 2004 to 2009, and in recent years, the return on assets (ROA), the total assets, and the ratio of non-current loans to allowance for loan losses. Finally, notice that there is a tendency for the number of important variables to increase over the testing period, with early models having only four or five variables, expanding to seven variables by 2011 (see last row of Figure 22).

The right panel of Figure 22 summarizes the consistency of variable contributions to the one-year models over time, both in terms of what percentage of the 13 yearly models each variable was included, but also its average place in the hierarchy of contributions if selected. As mentioned above, ROE was selected as the first variable using the BIC 100%

of the time, with the ratio of Liabilities to Assets, also chosen in all models, but having an average in the order of 2.1 owing to its third position in the 2010 model. The percentage of non-current loans and net operating income to assets are in 92% of the annual models at roughly fourth and fifth rank, whereas the fraction of earning assets is in 85% of the models, but when included has an average rank of 3.5. After those variables, contributions drop off rapidly, with the yield on earning assets in 46% of the models, but only at a rank of 5.8, followed by ROA and total assets at 28%. Finally, non-current loans to loan loss allowance is in 15% of the annual one-year models at a rank of fifth. Variables never included in any of the annual one-year models are: net loans to bank equity capital, annual default rate, ratio of assets 90 days past due to those 30-60 days past due, net interest income to earning assets, and net interest margin.

2 Year Model

Variables	Probability of Selection	Average Order if Selected
Non-Current Loans / Loans	100%	1.1
Net Loans / Bank Equity Capital	100%	3.6
ROE	92%	2.7
Yield on Earning Asset	85%	4.2
Earning Assets / Assets	85%	4.3
Assets 90 days Past Due / 30 - 89 Days Past Due	31%	5.8
Annual Default Rate	23%	5.3
Liabilities / Assets	23%	5.7
Assets	23%	5.7
Net Interest Income / Earning Assets	15%	8.0
Net Operation Income / Assets	0%	-
Net Interest Margin	0%	-
ROA	0%	-
Non-Current Loans / Loan Loss Allowance	0%	-

3 Year Model

Variables	Probability of Selection	Average Order if Selected
Net Loans / Bank Equity Capital	100%	1.8
Non-Current Loans / Loans	100%	2.1
Yield on Earning Asset	92%	4.3
Net Operation Income / Assets	69%	3.9
Liabilities / Assets	31%	3.0
Annual Default Rate	23%	2.0
Assets	23%	5.7
Assets 90 days Past Due / 30 - 89 Days Past Due	23%	6.0
ROE	15%	4.5
Net Interest Income / Earning Assets	15%	9.0
Net Interest Margin	8%	10.0
Non-Current Loans / Loan Loss Allowance	8%	11.0
Earning Assets / Assets	0%	-
ROA	0%	-

4 Year Model

Variables	Probability of Selection	Average Order if Selected
Yield on Earning Asset	77%	2.3
Non-Current Loans / Loans	69%	2.4
Net Operation Income / Assets	69%	2.7
Annual Default Rate	31%	2.0
Net Loans / Bank Equity Capital	31%	3.0
Liabilities / Assets	31%	4.0
Assets 90 days Past Due / 30 - 89 Days Past Due	23%	1.3
Assets	15%	4.0
ROA	8%	6.0
Net Interest Margin	8%	7.0
Net Interest Income / Earning Assets	8%	8.0
ROE	0%	-
Earning Assets / Assets	0%	-
Non-Current Loans / Loan Loss Allowance	0%	-

5 Year Model

Variables	Probability of Selection	Average Order if Selected
Yield on Earning Asset	54%	2.6
Assets	46%	4.0
Net Interest Margin	38%	3.6
Earning Assets / Assets	31%	1.0
Annual Default Rate	31%	2.5
Net Loans / Bank Equity Capital	23%	2.7
Assets 90 days Past Due / 30 - 89 Days Past Due	23%	3.0
Liabilities / Assets	23%	4.3
Net Interest Income / Earning Assets	23%	6.7
Non-Current Loans / Loans	15%	8.0
Net Operation Income / Assets	8%	2.0
ROA	8%	2.0
ROE	0%	-
Non-Current Loans / Loan Loss Allowance	0%	-

Figure 23. Summary of Variable Selection, Displaying Probability of Selection and Average Order if Selected for Annual Two-, Three-, Four- and Five-Year Bank Default Models

Figure 23 presents summaries of the consistency of variable contributions to the two-, three-, four- and five-year models over time, analogous to that shown for the one-year model at the right in Figure 22. (The lists of variables selected and their orders for the two- to five-year models analogous to that shown for the one-year model at the left in Figure 22 appear in Figure 24.) The figures reveal shifts in the importance of various predictive variables over time. For example, the ROE, most important for the one-year model, becomes successively unimportant for predicting default in later years, not even being included on any of the annual four- or five-year models. Liabilities to assets, also important at one year, declines immediately to around 30% at two-years and remains at about that frequency, but never of greater importance than a rank of third. Conversely, net loans to bank equity capital, not included in any one-year models, is in every model at two and three years, maintaining its contribution, albeit in lesser amounts, out to five years. Meanwhile, the yield on earning assets, only marginally important at one year, becomes more important at longer horizons, being one of the most important at three to four years.

2y Model													
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Variables	Order in Which Selected (1,2,...)												
Non-Current Loans / Loans	1	1	2	1	1	1	1	1	1	1	1	1	1
Net Loans / Bank Equity Capital	2	2	5	5	5	5	5	5	5	2	2	2	2
ROE		4	1	2	2	2	2	2	2	4	5	3	3
Yield on Earning Asset		5	3	3	3	3	3	3	3	5	6		9
Earning Assets / Assets	3	3	4	4	4	4	4	4	4	3			10
Assets 90 days Past Due / 30 - 89 Days Past Due										6	4	6	7
Annual Default Rate											3	7	6
Liabilities / Assets											9	4	4
Assets											7	5	5
Net Interest Income / Earning Assets											8		8
Net Operation Income / Assets													
Net Interest Margin													
ROA													
Non-Current Loans / Loan Loss Allowance													
Num of Vars Selected	3	5	5	5	5	5	5	5	5	6	9	7	10

3y Model													
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Variables	Order in Which Selected (1,2,...)												
Net Loans / Bank Equity Capital	2	2	2	2	2	2	2	2	2	1	2	2	1
Non-Current Loans / Loans	1	1	1	1	1	1	1	1	1	2	5	8	3
Yield on Earning Asset	3	3	3	4	4	4	3	3	3		8	6	8
Net Operation Income / Assets			4	3	3	3	4	4	4	3	7		
Liabilities / Assets										4	3	3	2
Annual Default Rate											1	1	4
Assets											6	5	6
Assets 90 days Past Due / 30 - 89 Days Past Due											4	7	7
ROE												4	5
Net Interest Income / Earning Assets												9	9
Net Interest Margin													10
Non-Current Loans / Loan Loss Allowance													11
Earning Assets / Assets													
ROA													
Num of Vars Selected	3	3	4	4	4	4	4	4	4	4	8	9	11

4y Model													
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Variables	Order in Which Selected (1,2,...)												
Yield on Earning Asset	1	2	2	2	3	2	2	2	2		5		
Non-Current Loans / Loans	2	1			1	1	1	1	1	5			9
Net Operation Income / Assets			1	1	2	3	3	3	3	2		6	
Annual Default Rate										1	4	2	1
Net Loans / Bank Equity Capital										3	2	3	4
Liabilities / Assets										4	3	4	5
Assets 90 days Past Due / 30 - 89 Days Past Due											1	1	2
Assets												5	3
ROA													6
Net Interest Margin													7
Net Interest Income / Earning Assets													8
ROE													
Earning Assets / Assets													
Non-Current Loans / Loan Loss Allowance													
Num of Vars Selected	2	2	2	2	3	3	3	3	3	5	5	6	9

5y Model													
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Variables	Order in Which Selected (1,2,...)												
Yield on Earning Asset						1	3	1	2		1	5	5
Assets						2	2	2	1			9	8
Net Interest Margin	1	1									4	6	6
Earning Assets / Assets				1	1		1			1			
Annual Default Rate								3			3	2	2
Net Loans / Bank Equity Capital											2	3	3
Assets 90 days Past Due / 30 - 89 Days Past Due											7	1	1
Liabilities / Assets											5	4	4
Net Interest Income / Earning Assets											6	7	7
Non-Current Loans / Loans											8	8	
Net Operation Income / Assets	2												
ROA		2											
ROE													
Non-Current Loans / Loan Loss Allowance													
Num of Vars Selected	2	2	0	1	1	2	3	3	2	1	8	9	8

Figure 24. Left: Potential Model Variables and Their Order of Selection for the Two Through Five-Year Models for 2000 to 2012

A more detailed summary of the changes in variables as the annual prediction horizon increases from one to five years appears in Figure 25. The most important variables for each annual prediction horizon are listed followed by a description of the changes that occurred from the previous years' model. For example, in predicting default from one to two years, ROE has become less important and the ratio of liabilities to assets is no longer in the model, whereas the fraction of non-current loans has become important along with the ratio of net loans to bank equity capital. In moving from the two- to three-year horizon, ROE is no longer in the model, whereas net loans to bank equity capital, which entered the model in at two years, is now most important. When predicting default between three and four years, the value of non-current loans is no longer in the model, whereas the current yield on earning assets has become most important. Finally, at five years, the current yield on earning assets remains most important along with net interest margin, with the fraction of current loans and operating income to assets no longer included. In general, quality of assets is most important in near-term default predictions, giving way to operating income and the yield on earning assets as the important determinants of default at longer horizons. It is important to remember that our models are for marginal annual defaults. Thus, variables important for modeling defaults in early years remain important for cumulative default prediction, just not for predicting marginal defaults pending survival to later years.

1-Year Model

- Most Important Variables:
  - Return on Equity
  - Liabilities/Assets
  - Percent of Non-Performing Loans
- Number of variables tends to increase with length of training period
  - Early models have 4-5 variables; later models have 6-7 variables

2-Year Model

- Most Important Variables:
  - Fraction of Non-Current Loan
  - Net Loans / Bank Equity Capital
  - Return on Equity
  - Yield on Earning Assets
  - Earning Assets/Assets
- Biggest Changes from 1- to 2-Year Models
  - Net Operating Income / Assets has become relatively unimportant at two years
  - Liabilities / Assets has lost importance

3-Year Model

- Most Important Variables:
  - Net Loans / Bank Equity Capital
  - Non-Current Loans / Loans
  - Net Operating Income / Assets
- Biggest Changes from 1- and 2-Year Models to 3-Year Model
  - Earning Assets / Assets is insignificant at 3 years
  - ROE has continued to decrease in importance between 1 and 3 years

4-Year Model

- Most Important Variables:
  - Yield on Earning Assets
  - Net Current Loans / Loans
  - Net Operating Income / Assets
- Biggest Changes from 1- through 3-Year Models to 4-Year model
  - Non-Current Loans is insignificant at 4 years
  - Net Interest Margin has not been important for years 1-3

5-Year Model

- Most Important Variables:
  - Yield on Earning Assets
  - Assets
  - Net Interest Margin
  - Earning Assets / Assets
- The consistency of variable selection tends to decrease as prediction horizon increases

Figure 25. Description and Comparison of Important Variables for Bank Defaults at One- to Five-Year Horizons

A quantitative analysis of the variable contributions over time is presented in Figure 26. The left table in the figure shows the likelihood of each variable being selected for the default models by horizon in each year over the period from 2000 to 2012. The table at the right displays the average order of selection if the variable was included in the model at the listed horizons. Consider first, the probabilities of variable selection. Those probabilities have been color coded for convenience, with variables included in 76% to 100% of the annual models for a given horizon coded in red, those included between 11% and 75% in green, and those in 10% or less in blue. Although ROE and the ratio of liabilities to assets are included in all one-year models, their contributions drop off rapidly at two to five years. Also, performance on earning assets (earning assets to assets) is important at one- and two-year horizons, but is not included in three- and four-year models, with only moderate contributions to five-year models. The color coding in Figure 26 helps to reveal some features not easily distinguished in the data. For example, the most consistently important variable is the percentage of non-current loans, included in all annual models from one to three years and dropping off to 69% at four years and 15% at five years. Also, asset size, while not in all annual models is in 15% and 46% of models in all years. For two- to four-year horizons, the yield on earning assets and net loans to bank equity capital become important, while having less influence in one- and five-year models. Notably, the five-year models appear to have the most diversity of variable contributions with no model in more than 54% of annual models. Finally, net interest margin appears relatively unimportant, except at the five-year horizon, with the ratio of non-current loans to loan loss allowance only included in a small fraction of models at one- and three-year horizons.

Variables	Probability of Selection					Average Order if Selected				
	1y	2y	3y	4y	5y	1y	2y	3y	4y	5y
ROE	100%	92%	15%	0%	0%	1.0	2.7	4.5	-	-
Liabilities / Assets	100%	23%	31%	31%	23%	2.1	5.7	3.0	4.0	4.3
Non-Current Loans / Loans	92%	100%	100%	69%	15%	3.9	1.1	2.1	2.4	8.0
Net Operation Income / Assets	92%	0%	69%	69%	8%	4.9	-	3.9	2.7	2.0
Earning Assets / Assets	85%	85%	0%	0%	31%	3.5	4.3	-	-	1.0
Yield on Earning Asset	46%	85%	92%	77%	54%	5.8	4.2	4.3	2.3	2.6
ROA	23%	0%	0%	8%	8%	4.0	-	-	6.0	2.0
Assets	23%	23%	23%	15%	46%	5.7	5.7	5.7	4.0	4.0
Non-Current Loans / Loan Loss Allowance	15%	0%	8%	0%	0%	5.0	-	11.0	-	-
Net Loans / Bank Equity Capital	0%	100%	100%	31%	23%	-	3.6	1.8	3.0	2.7
Annual Default Rate	0%	23%	23%	31%	31%	-	5.3	2.0	2.0	2.5
Assets 90 days Past Due / 30 - 89 Days Past Due	0%	31%	23%	23%	23%	-	5.8	6.0	1.3	3.0
Net Interest Income / Earning Assets	0%	15%	15%	8%	23%	-	8.0	9.0	8.0	6.7
Net Interest Margin	0%	0%	8%	8%	38%	-	-	10.0	7.0	3.6

Figure 26. Probabilities of Variable Selection by Model Horizon (Left) and Average Order of Inclusion If Selected at Given Horizon (Right) For Annual Models Over the Period from 2000-2012

The average order of selection of each variable for each model horizon appears at the right in Figure 26. Variables selected early in the process (averaging from 1 to 2.5) are coded in red, those in the middle set, averaging between 2.5 and 5.0 are in green, and those selected above fifth on average in blue. The table shows the importance of ROE in early year models; ROE was selected first in all one-year models from 2000 to 2012, with non-current liabilities second in all but one year. Again, for mid-year horizons, net operating income becomes selected in most models, averaging between 1.1 and 2.4 in selection order. Although the annual default rate is never included in more than 31% of models at any horizon, when selected at four- and five-year horizons, its average order is second. The table also confirms the relative unimportance of net interest income to earning assets and net interest margin.

The data at the right in Figure 26 seem to suggest that a variable's probability of being selected is related to the order in which is it selected. This is confirmed, at least in general, by Figure 27, which demonstrates that the average order of selection (first to last) increases with the probability of selection. That is, frequently selected variables tend to be selected earliest and vice versa. Examples of frequently selected variables with early selection include ROE, the ratio of liabilities to assets, and percentage of non-current loans. Conversely, rarely selected variables such as net interest income to earning assets and net interest margin are typically near the last ones to be selected in the few models in which they are present.

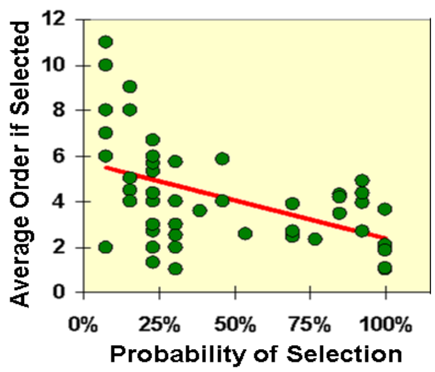


Figure 27. Relationship between Variables' Probability of Selection and Order of Importance

I conclude my analysis of variable contributions with an analysis of the consistency of the signs of the variables input into each model. That is, I measure the extent to which a variable, if selected maintains the same sign in different years and tenors. For example, variables having positive signs are indicative of higher default risk, whereas negative inputs signal lower risk of default. Figure 28 shows for each variable the likelihood that it is positive if selected. Again, color coding is used to highlight variables; variables having 67% or greater likelihood of being positive highlighted in red (for higher risk), those between 0% and 33% positive in green (for lower risk), and gray for those in between. If a variable is not included in a model (e.g., for ROE in year four and year five models) the corresponding cell is blank. Clearly, almost all variables, when selected for a model in a given year are of the same sign. That is, nearly all cells with values are either 0% or 100%. Exceptions to this include assets at the five-year horizon and net operating income to assets at one year. Although some variables are of the same sign when included in models at all horizons (e.g., ROE, annual default rate, and net interest income to earning assets), most variables change sign when included at different horizons. However, excluding their inputs to the five-year model, all variables except percent of non-current loans are of the same sign if included in models at shorter horizons.

Variables	Prob Positive if Selected					Average Value if Selected				
	1y	2y	3y	4y	5y	1y	2y	3y	4y	5y
ROE	0%	0%	0%			-0.6	-0.6	-0.5		
Liabilities / Assets	100%	0%	0%	0%	100%	1.0	-1.0	-1.2	-1.4	1.7
Non-Current Loans / Loans	100%	100%	100%	100%	0%	0.6	1.1	0.9	0.6	-0.9
Net Operation Income / Assets	17%		0%	0%	100%	-0.6		-0.7	-0.7	1.4
Earning Assets / Assets	0%	0%			0%	-0.5	-0.7			-8.1
Yield on Earning Asset	100%	100%	100%	90%	0%	0.4	0.7	0.7	0.7	-3.7
ROA	0%	-	-	0%	100%	-1.5			-0.3	1.2
Assets	100%	100%	100%	100%	50%	0.3	0.5	0.5	0.5	-1.4
Non-Current Loans / Loan Loss Allowance	100%		0%			0.6		-0.4		
Net Loans / Bank Equity Capital		100%	100%	100%	0%		0.7	0.8	1.6	-0.5
Annual Default Rate		0%	0%	0%	25%		-0.2	-0.5	-0.3	0.0
Assets 90 days Past Due / 30 - 89 Days Past Due		0%	0%	0%	0%		-0.6	-0.6	-0.5	-4.4
Net Interest Income / Earning Assets		0%	0%	0%	67%		-0.4	-0.6	-1.1	0.9
Net Interest Margin			100%	100%	0%			0.5	1.2	-4.0

Figure 28. Analysis of Variable Consistency by Model Tenor: Left: Probabilities of Variables Having Positive Signs (Higher Default Risk; Red) If Selected and Right: Average Values of Variable Contributions If Selected

A clue to the changes in signs of variables at five years is the large average coefficients, both positive and negative assigned to variables in the five-year models. Many of those values (e.g., fraction of earning assets, assets 90 days past due, net interest margin, yield on earning assets are four to eight times the average values of variables in models at shorter tenors. The results in Figure 26 show that the consistency of variables selected for the models decrease with tenor and these instabilities likely underlie the large values selected for input variables when chosen for fifth-year default predictions.

7. Conclusion

In this study, I develop a dynamic measure to overcome limitations of the Merton-type structural models in predicting default probabilities for financial firms. I build and test adaptive statistical models to estimate default probabilities for U.S. banks. As described in detail, the models are logistic regression whose input variables are selected and calibrated based on their past effectiveness at predicting bank failures. Selection of variables in the model and their weights were updated yearly using a "walk-forward" procedure.

I presented a detailed analysis of the contributions of financial variables to model outputs by year (2000-2012) and tenor (1-5 years ahead). For a given prediction horizon, I find great consistency among variables selected for each annual model over the period from 2000 to 2012. For predictions of marginal defaults in years one and two, return on equity and percentage of non-performing loans are major determinants of default. Those variables give way in importance at intermediate tenors (i.e., three- to four-years) to current yield on earning assets and net loans to bank equity capital. Finally, at the five-year horizon, yield on earning assets, asset size, and net interest margin become most important predictors of default.

I also analyzed the consistency of the order of variable selection as well as the signs of the variables as they reflect increases and decreases in credit quality. I find a rough, but positive, relationship between the probability of a variable being selected in each annual model and its average importance in the predictive selection hierarchy. Also, I find that variables, when selected for a given predictive tenor, are nearly always of the same sign across the years of annual model development. However, I do find that many variables change sign as the year of marginal default prediction changes, but this mostly occurs for a single year; predicting defaults between years four and five.

I further derived estimates of physical default probabilities from risk scores from the bank PD model. These PDs can then be used to estimate expected losses from default on bank portfolios. Because most models offer little benefit over agency ratings for already low-rated credits, I examined the performance of the bank PD model relative to Kroll agency ratings for credits rated above single-B-minus and for those rated above single-C-minus. I find that the predictive power of agency ratings drops off dramatically as credit quality of the scoring sample increases, with much less deterioration in default prediction using the bank PD model.

The model predicts defaults at annual horizons from one to five years. Performance of the models at discriminating between defaults and non-defaults is evaluated for horizons of one to five years using a sequence of annual walk-forward out-of-sample tests from 1992 to 2012. I also measure the ability of the models to predict absolute default rates from one to five years and, except for underestimating the high bank default rates during the financial crisis, the models performed well at estimating the annual bank default rates. In general, the models perform favorably at predicting defaults, with a 97% accuracy ratio (AR) at one year prior to default, and decreasing, but still above-chance predictive power out to five years. The models are designed to be updated on an annual basis, but updated financials for inputs to the model are available from the FDIC on a quarterly basis. Not only do my results provide evidence of the advantage of the adaptive statistical modeling approach over agency ratings, but also they provide insight to the short- and long-term determinants of bank failure.

References

Altman, E. (1968). Financial Ratios, Discriminant Analysis, and the Prediction of Corporate Bankruptcy, *Journal of Finance, 23,* 589-609. http://dx.doi.org/10.1111/j.1540-6261.1968.tb00843.x

Beaver, W. H. (1966). Financial Ratios as Predictors of Failure, *Journal of Accounting Research, 4,* 71-111. http://dx.doi.org/ 10.2307/2490171

Department of the Treasury, Office of the Comptroller of the Currency (2012). 12 CFR Parts 1, 5, 16, 28, and 160 [Docket ID OCC–2012–0005] RIN 1557–AD36, Alternatives to the Use of External Credit Ratings in the Regulations of the OCC, Federal Register, *77*(114). http://www.gpo.gov/fdsys/granule/FR-2012-06-13/2012-14169

Eom, Y., Helwege, J., & Huang, J. (2003). Structural Models of Corporate Bond Pricing: An Empirical Analysis, Forthcoming: Review of Financial Studies.

Huang, J., & Huang, M. (2000). How Much of the Corporate-Treasury Yield Spread is Due to Credit Risk? A New Calibration Approach, Working Paper, Cornell University.

Kealhofer, S. (1999). Credit Risk and Risk Management, AIMR Conference Proceedings, 80-94. http://dx.doi.org/10.4172/2151-6219.1000126

Moody's Investors Service (2010). Corporate Default and Recovery Rates, 1920-2009. http://dx.doi.org/ 10.3905/jfi.2011.21.2.015

Schwarz, G. E. (1978). Estimating the Dimension of a Model, *Annals of Statistics, 6*(2), 461–464.

Sobehart, J., & Stein, R. (2000). Moody's Public Firm Risk Model: A Hybrid Approach to Modeling Short Term Default Risk, Moody's Investors Service, Global Credit Research.

Sobehart, J., & Keenan, S. (2002). Hybrid Contingent Claims Models: A Practical Approach to Modeling Default Risk.

Sobehart, J., & Keenan, S. (2003). Hybrid Probability of Default Models: A Practical Approach to Modeling Default Risk, *Quantitative Credit Analyst, 3,* 5-29.

Vasicek, O. (1984). Credit Valuation, KMV White Paper, KMV LLC.