JULY 2002



MOODY'S
ANALYTICS

# RATING METHODOLOGY

# Moody's RiskCalc™ Model for Privately-Held U.S. Banks


### Authors

Ahmet E. Kocagil
Alexander Reyngold
Roger M. Stein
Eduardo Ibarra

### Contact Us

To learn more, please contact your Moody's KMV client representative, visit us online at www.moodyskmv.com, contact Moody's KMV via e-mail at info@mkmv.com, or call us at:

Ahmet E. Kocagil 1.212.553.1653

André Salaam

Alternatively, you may contact our customer service team:

| | |
|---|---|
| Americas | +1.212.553.1653 |
| Europe | +44.20.7772.5454 |
| Asia-Pacific | +85.2.2916.1121 |
| Japan | +81.3.5408.4100 |


## Rating Methodology

This report documents RiskCalc for U.S. Banks, Moody's model for estimating the probability of default (PD) for privately-held U.S. banks, thrifts, and bank holding companies. RiskCalc for U.S. Banks is a robust and validated model that produces one- and five-year PDs. It predicts separate PDs for bank holding companies and bank and thrift subsidiaries.

RiskCalc is a statistical model that estimates PD based on financial statement data. The model has been tested using rigorous out-of-sample [and out of time] techniques to prevent overfitting. The model is based on data from approximately 7,000 U.S. banks, thrifts, and bank holding companies including over 400 failures over the last two decades.

We believe RiskCalc for U.S. Banks is an important addition to the RiskCalc network of default prediction models, which now includes country-specific models for private companies in the U.S., Canada, Mexico, United Kingdom, Germany, France, Spain, Portugal, Netherlands, Belgium, Japan, and Australia. In addition, Moody's KMV provides global coverage of publicly traded industrial and financial firms. We feel that RiskCalc for U.S. Banks will be a meaningful addition to the practice of credit risk management and a step forward in answering the call for rigor that the BIS has outlined in their recently proposed Basel Capital Accord.

# CONTENTS

## Highlights

1. We describe the methodology followed in estimating Moody's RiskCalc™ model for U.S. Banks, which is estimated using actual bank failure data on U.S. banks, thrifts, and bank holding company defaults.[1] We describe the factors in the model, the modeling approach, and the model's accuracy.

2. We find that RiskCalc for U.S. Banks performs better at predicting bank failures than other publicly available models. Specifically, RiskCalc exhibits higher power over both one- and five-year horizons when tested using out-of-sample [and out-of-time] techniques.

3. We find this performance to be robust not only over different historical periods, but also across banks, thrifts, and bank holding companies. We distinguish between BHCs, commercial banks, savings banks, and savings institutions. These distinctions are based on what type of regulatory report an institution files. BHCs file the Y-9 report with the Federal Reserve Board. Commercial Banks and Savings Banks file the FDIC Call Report (the further distinction between the two is determined by an institution's charter type). Savings Institutions file the OTS Thrift Financial Report.

4. RiskCalc for U.S. Banks is based on the credit experience of over 7000 U.S. banks, thrifts, and bank holding companies including over 400 defaults over the last two decades.

This document is a self-contained description of the development and validation of RiskCalc for privately-held U.S. Banks.

## 1. Overview Of RiskCalc For U.S. Banks

### 1.1 MODEL FACTORS

Recently, regulatory bodies have focused more closely on probability of default (PD) analysis. The proposed *New Basel Capital Accord* (Basel, 2001) addresses the issue explicitly (Basel, § 173). Accordingly, banks would have the option of using conservative pre-defined PD measures under the so-called foundation approach, but if they wish to qualify for the advanced approach:

> *…A bank must estimate a PD for each of its internal PD grades…Each estimate of PD must be grounded in historical experience and empirical evidence. At the same time, these estimates must be forward looking…PD estimates that are based purely on subjective or judgmental consideration and not grounded in historical experi- ence and data will be rejected by supervisors. (Basel, § 336 & 337).*

RiskCalc is meant to address this growing need in financial markets and provide an objective, quantitative default prediction tool for credits.

RiskCalc for U.S. Banks incorporates market and financial statement information based on empirical data and statistical evidence. The selection of financial statement variables is informed by the collective experience of Moody's analysts. But as different experts prefer different ratios, the number of candidate ratios for developing a bank default model is quite large. Our preference is for the smallest number of inputs and the simplest functional form. Clearly, it is not feasible to include every possible financial ratio. The result of an exhaustive approach would likely make the model highly susceptible to overfitting, i.e. estimating a model that may work well in-sample but would have inferior out-of-sample power.

1For consistency with the existent RiskCalc network of models, in this document we use the terms 'default' and 'probability of default (PD)'to indicate bank failures and probabilities associated with these events, respectively.

Therefore, based on the findings of previous studies focusing on empirical bank default modeling literature and consultations with banking sector analysts, we constrained our approach to six major categories of variables specific to the banking industry:

» capital,

» asset quality,

» concentration,

» liquidity,

» profitability, and

» growth.

Within these categories we selected those factors that demonstrated high power on a univariate basis and low correlation with other factors. The result was a set of ten ratios that form the basis RiskCalc bank model (See Section 2.2 for details). In addition, we include a macroeconomic indicator variable, the monthly deviation of the speculative grade default rate from its 24-month trailing average, to signal the state of the credit cycle.

We believe that RiskCalc for U.S. Banks is a meaningful addition to the practice of credit risk management and a step forward in answering the call for rigor that the BIS has outlined in their recently proposed Basel Capital Accord.

### 1.2 EMPIRICAL RESULTS

We performed extensive testing of the model to ensure that the performance is stable over time and that the model has not overfit the data. To validate the performance of RiskCalc for U.S. Banks, we have adopted two approaches that have been adopted by Moody's to validate default prediction models: *k-fold* analysis and *walk-forward* analysis. *K-fold* analysis tests the robustness of the model vis-à-vis different subsamples and ascertains that no particular data period is "driving" the model. The *walk-forward* validation, on the other hand is a robust means of ensuring that models have not been "overfit" to the data, that the future performance can be well understood, and that the modeling approach, as well as any individual model produced by it, is robust through time and credit cycles. This approach involves testing models on data not used to develop them and, when possible, on data associated with future time periods. This ensures that the modeling approach is able to capture changes in the credit cycle. The methodology is described more fully in (Sobehart, Keenan, and Stein, 2000).

In brief, we find that RiskCalc model for privately-held U.S. banks is both powerful in discriminating failing from non-failing institutions over one- and five-year horizons and exhibits consistent and high power across banks, thrifts, and bank holding companies.

### 1.3 DATA USED IN THE DEVELOPMENT OF RISKCALC FOR U.S. BANKS

The RiskCalc model was constructed using information on 17,673 unique banks, thrifts, and bank holding companies. A total of 161,034 observations from 1986 to 1999 were used.

We use the FDIC definition of bank failure in our modeling exercise[2] and, similar to Wheelock and Wilson (2000), also include FDIC financial assistance, assisted mergers, and payoffs as defaults. As of June 2000, our database contained more than 2,400 public and private defaults in the banking industry.

In order to avoid any systematic bias in our modeling efforts, e.g. the crash in 1980s, we excluded thrift failures from our model estimation sample (although we tested the model using thrift defaults and we report the results of those tests here).

2 For details, see http://www.fdic.gov.

We further corrected for failures of the same institution in several locations. After this "double-counting" adjustment we had a reduced sample of 413 private bank failures suitable for modeling. The resulting private bank defaults were fairly evenly distributed in time and capture the "default wave" of late 1980s and early 1990s. Moreover, the data also covers the post-default wave period of the mid-1990s to 2000, so the model coverage is relatively unbiased and reflects an entire cycle in terms of historical data.

Financial statement and other data were collected from several commercial sources as well as from Moody's proprietary databases.

## 2. Modeling Methodology

### Differences In Industry And Model Structure

The banking industry and thus the financial sector in general exhibit several different characteristics relative to the non-financial universe.

The first and most general distinction between the two universes is the typical cause of default. While the cause of default can in general be characterized as either business risk or financial risk for corporate firms, bank defaults tend to be caused much more often by financial risk. In other words, while a manufacturing firm engages in financial transactions to support its operations, a bank performs the same activity as its core business. This distinction suggests that bank default prediction models and industrial default models ought to focus on different factors. Thus, for banks, we consider factors such as portfolio structure, asset quality, and so forth in addition to the more standard factors of default.

By definition, the banking industry consists of several types of institutions that are specialized in their activities, and scope, i.e. commercial banks, savings banks, savings associations, etc. Thus, a model must balance the specificity and weights of its factors in encompassing the entire universe.

Furthermore, most large corporate firms trade individually in order to increase capital market accessibility. In contrast, a notable portion of the largest U.S. banks do not trade in equity markets, per se but shares are trading for their parent institutions, typically a bank holding company (BHC). In addition, unlike diversified holding companies in the corporate universe, it would not be atypical for bank holding companies to have an overall riskiness level that is higher than its subsidiaries. Given these observed differences in the banking industry, it is clear that one can not apply the standard RiskCalc corporate firm model in evaluation institutions in this industry.

### 2.1 THE MODELING PROCESS

RiskCalc for U.S. Banks is an empirical model, in that although it is informed by the collective experience of Moody's analysts, it is based on empirical data and statistical evidence. The first step in empirical model development is the selection of input variables. With a substantial number of ratios preferred by expert financial analysts, the number of candidate ratios is quite large. Clearly, it is not feasible to consider each and every possible definition in all categories of financial ratios. Furthermore, the result of an exhaustive step-forward approach would make the model highly susceptible to overfitting. Our bias is towards the simplest functional form and the smallest number of inputs.

Therefore, based on the findings of previous studies focusing on empirical bank default modeling literature[3] and discussions with Moody's analysts, we constrained our approach to the six major categories: capital, asset quality, concentration, liquidity, profitability and growth. Within these categories we selected those factors that demonstrated high power on a univariate basis and low correlation with other factors. The result is ten ratios that form the basis of the RiskCalc banking model. In addition, when evaluating banks that are subsidiaries of publicly traded bank holding companies ('hybrids'), we also use equity market information of the 'parent' institution.[4]

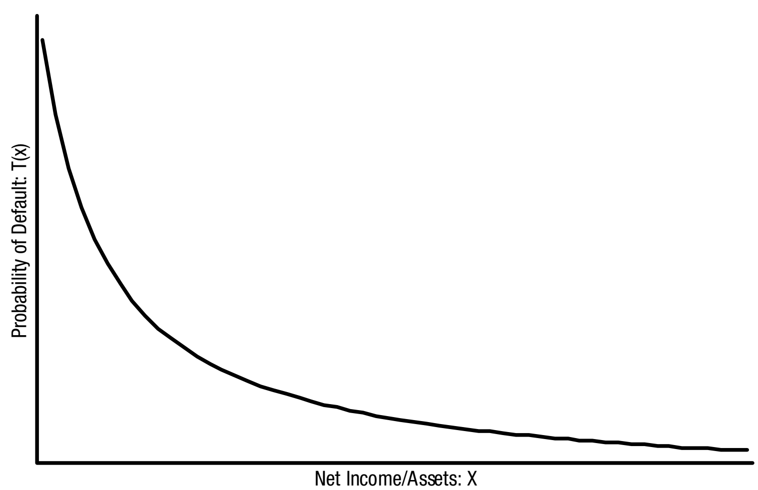
Our modeling approach, which is similar to the previous RiskCalc models,[5] can be briefly summarized in three steps: transformation, modeling, and mapping.

### 2.1.1 Transformation

The first step involves replacing each financial ratio with an estimation of its corresponding univariate default probability over 5 years. That is, each ratio, such as net income/assets, is related to a default proba- bility; instead of using the raw (and often noisy) input ratio, we use the default probability corresponding to this ratio. We call this process "mini-modeling." This captures much of the nonlinearity of the problem, normalizes the inputs to a common scale, controls for outliers, and helps monitoring the marginal effect within the model simply by observing the univariate default prediction.

**Figure 1** (below) shows an example of a transformation function used within the model. In this case, it approximately reflects the behavior of Net Income/Assets. The transformation turns the ratio, which fluc- tuates between -0.25 and 0.50, into probability space so that it can be mapped to values between zero and, for example, 0.10%. Note the variable's nonlinear relationship to default. In the example below, the steep slope of the transformed variable on the left side implies a significant risk difference between the least prof- itable and the moderately profitable companies, while in turn there is little difference between companies with very high and moderate profitability.

FIGURE 1: Typical Transformation Function




*Transformations capture nonlinear relationships and illustrate ratios' marginal effects*

3 Some examples of previous research on bank default modeling are (1) Peek, Joe, and Eric S. Rosengren. 1997. How Well-Capitalized are Well-Capi- talized Banks? New England Economic Review, September-October: 41-50. (2) Meyer, Paul A. and Howard W. Pifer. 1970. Prediction of Bank Fail- ures. Journal of Finance. September 1970: V25, number 4. (3) Thompson, James B. 1991. Predicting Bank Failures in the 1980s. Federal Reserve Bank of Cleveland Economic Review. First Quarter: 9-20. (4) Gilbert, R. Alton, Andrew P. Meyer and Mark D. Vaughan. 1999. The Role of Supervisory Screens and Econometric Models in Off-Site Surveillance. Federal Reserve Bank of St. Louis, November/December 1999: 31-56. (5) Estrella, Arturo, Sangkyun Park, and Stavros Peristiani. 2000. Capital Ratios as Predictors of Bank Failure. Federal Reserve Bank of New York Economic Policy Review. July. For a detailed discussion on the previous literature please see Appendix C.

4 We use this information to calculate the distance to default based on a structural model motivated by Merton (1974), which we also transform and enter into the logistic regression along with the other variables in the model in evaluating hybrid institutions.

5 For a detailed discussion of this approach, please refer to: RiskCalc for Private Companies II (2000).

### 2.1.2 Modeling

The second step in the RiskCalc modeling process involves combining the transformed inputs in a multivariate model so that weights can be assigned to the multivariate model. Similar to previous RiskCalc Private firm models, the selected variables are regressed on a dummy variable, which flags actual default events in a probit model framework. The resulting model generates a financial score (FS), which is based on company-specific financial statement information and a macro variable depicting credit quality in the market environment.

We use the univariate transformations (mini-models) discussed in the previous section as inputs to a binary choice model that predicts default. RiskCalc for U.S. Banks is estimated using a probit model, which uses the normal or Gaussian cumulative distribution function, specifically:

$$y = Prob\ (default \mid x; B, \sigma) = F(\beta' T(x)) = \int_{-\infty}^{\beta' T(x)/\sigma} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{t^2}{2}\right) dt$$

The advantage of the probit model, as opposed to, say, ordinary least squares, is that it specifically accounts for the binary output (i.e., 0 or 1), which characterizes the "default/no default" nature of the default prediction problem. The resulting empirically estimated model is a generalized linear model in that it is a nonlinear[6] function of a linear model: $y = \Phi\ (f\ (x, B))$, where the linear part is simply

$$f\ (x, B) = B_0 + T_1\ (x_1)\ B_1 + T_2\ (x_2)\ B_2 + \ldots + T_{10}\ (x_{10})\ B_{10}$$

the $T(x_i)$ terms stand for the transformations of $x_i$.

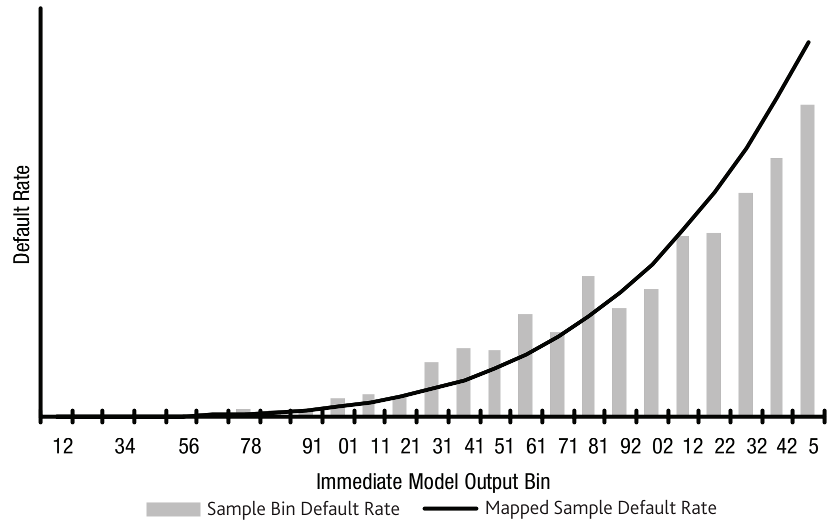
### 2.1.3 Mapping

The final stage of the modeling process, mapping, is similar to the transformation of ratios. We take the output of the multivariate probit model and use the empirical database to map the output into the sample default probability. This is done because invariably the output from the probit model tends to over- or underestimate the true probability within sample. It is a common problem in applied probit or logit prediction, and relatively straightforward to correct.

Figure 2 below shows how we take the output of the model and map it to sample default probabilities. We estimate the relation between model output and sample default probability using smoothing algorithms identically to the manner in which we smooth our input transformations for mini-modeling.

6 Specifically, a sigmoidal ("s-shaped") function.

FIGURE 2: Mapping To Sample Default Probabilities




*Model output is calibrated to sample default probabilities via a smoothing algorithm*

In sum, the transformations normalize the input ratios and capture their nonlinearity in a transparent way. Applying a binary choice model to these transformed inputs creates outputs that must be mapped to default probabilities the same way that mortality tables are assigned to people of a certain age: based on historical data. The final adjustment is the adjustment for the top-down default rate for the entire sample, which is usually arrived at outside the model's three main stages (i.e., transformation, modeling, and mapping).

We adjust the *sample* default probability to our projected central tendency rate of default for the population. This is done because the sample may yield a biased default probability estimate. In order to implement this adjustment, we multiply the produced PD scores by a constant to impose our projected 1- and 5-year central tendency default rates, which are 0.70% and 2.50% respectively. These rates reflect the empirical default rate in the sector.[7]

The aforementioned adjustment can be shown simply as:

$$PD_{Population} = PD_{Sample} \ x \ \frac{central\ tendency\ rate}{sample\ failure\ rate}$$

## 2.2 MODEL FACTORS

The modeling of unlisted U.S. banks follows the same analytic steps as outlined above: first, we identify an extensive list of potential candidate variables. Next, for each input, x, we examine its relation to defaults at 1 and 5-year horizons. The obtained relationships, T(x), are then combined using a probit regression with the bank default events as the dependent variable.[8]

7 The average annual default rate for the 1985-2000 period for the overall banking industry is about 1.18%. The corresponding figure for the same universe net of OTS-supervised savings institutions is about 0.77%. Similarly, the same annual averages are 0.82% and 0.49% for the period 1990-2000, respectively. Our central tendency figure is based on triangularized use of these figures.

8 As a standard statistical check, we examined the correlation matrix of the model variables in both transformed and non-transformed space, and verified that there are no statistical problems with it.

### 2.2.1 Overview Of Model Factors

The variables in the model cover *profitability* (Net Income/Assets), *leverage* (Equity/Assets), *growth* (Liabilities Growth), *efficiency* (Net Interest Margin), *loan portfolio composition and concentration* (C&I Loans/Assets, Construction Loans/Assets, Commercial Real Estate Loans/Assets), holdings of *risk-free securities* (Govern- ment Securities/Assets), and charge-offs by loan types (Commercial Charge-Offs/ Assets, Installment Charge-Offs/Assets) to measure *asset quality*. We also include a macro indicator of the credit environment in the form of a trailing speculative grade default rate. In evaluation of banks with publicly traded parent we also use distance to default[9] and an estimate of the long-term credit quality of the parent as model inputs.[10]

In the following subsection we discuss each of the model inputs by category. **Table 1** presents a categorical representation of the model's inputs.

TABLE 1: Model Inputs By Category

| |
|---|
| **Capital** |
| Equity Capital/Assets |
| **Asset Quality** Commercial |
| Charge-Offs/Assets Installment |
| Charge-Offs/Assets |
| **Concentration** |
| Commercial Real Estate Loans/Assets |
| Construction Loans/Assets |
| C&I Loans/Assets |
| **Liquidity** |
| Government Securities/Assets |
| **Profitability** |
| Net Interest Margin Net Income/Assets |
| **Growth** |
| Liabilities Growth |

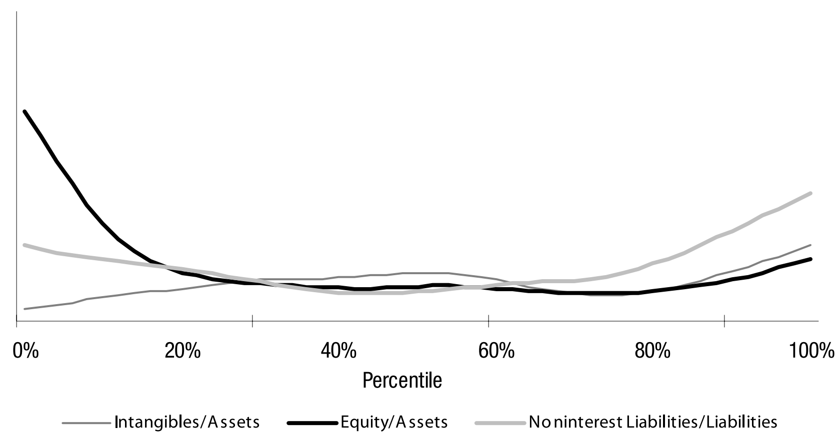
### 2.2.2 Capital

The importance of Tier 1 capital in credit risk modeling has been emphasized and put into regulatory forefront by the Basel Committee in the form of capital requirement ratios. Nevertheless, empirically, this ratio does not have a long enough history at this point to allow its use in modeling. In the United States, minimum capital ratios have been requested since 1981, and the Basel Accord has applied its capital ratio only since 1988. Thus, we note that the variable has not been reported for a long enough period to evaluate its predictive power in combination with other variables. Furthermore, the Basel capital requirement framework is undergoing conceptual changes with the purpose of determining more risk driven capital requirements.

As an alternative, we chose a different measure of capital in the form of the ratio Equity/Total Assets. As Estrella, Park, and Peristiani (2000) illustrate, simple leverage ratios such as this one predict bank default as well as much more complex risk-weighted ratios over one or two year horizons. **Figure 3** shows the relationship to default of the various leverage ratios we considered. The figure demonstrates that the Equity/Assets is the most informative of the candidate leverage measures in our model.

9 Distance to Default is the distance (in standard deviations of asset volatility) between the value of assets and the par value of liabilities. This variable is estimated base on a structural model that is motivated by Merton (1974).
10 Long-term credit quality is estimated based on asset size, leverage and volatility of parent's equity price.

FIGURE 3: 5-Year Default Probabilities For Capital Ratios

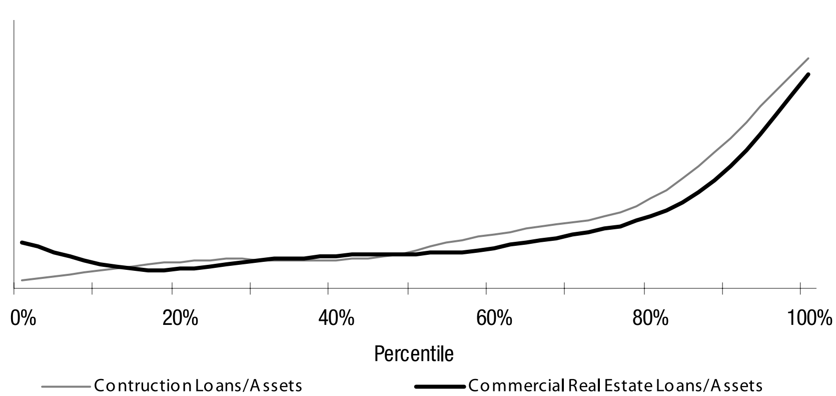



### 2.2.3 Asset Quality

To determine the asset quality of a bank, we looked at two key dimensions: asset concentration and credit quality. Generally, concentration is defined as the presence of a significantly large volume of economically related assets that an institution has advanced or committed to one entity or affiliated group. Concentration may in the aggregate present a substantial risk to the solvency and soundness of the institution. Additionally, concentration can occur due to geographic factors: this may occur when financial institutions are concen- trated in a particular region and are limited with respect to their ability to diversify.

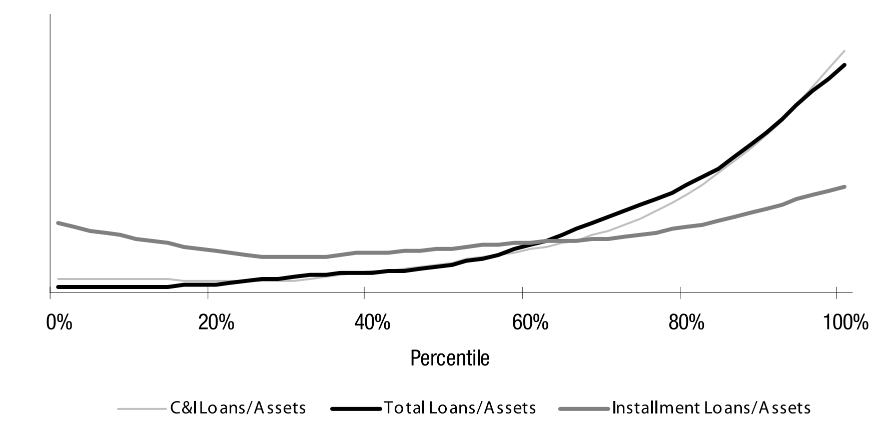
For our purposes, we are particularly interested in the composition of a bank's loan portfolio. The FDIC reported that some banks that failed from 1980 to 1994 lacked diversification. The incidence of default was particularly high in states characterized by a specific diversification concern: real estate downturns (e.g., California). Therefore, a high concentration in real estate or highly cyclical industries, such as construction, tends to increase the risk profile of a bank's portfolio. Therefore, we consider real estate concentration a particularly important measure of risk and we use real estate loans over assets and construction loans over total assets as key ratios. As shown in **Figure 4**, real estate exposure, as measured by these ratios, clearly adds to credit risk.

FIGURE 4: 5-Year Default Probabilities For Concentration Ratios



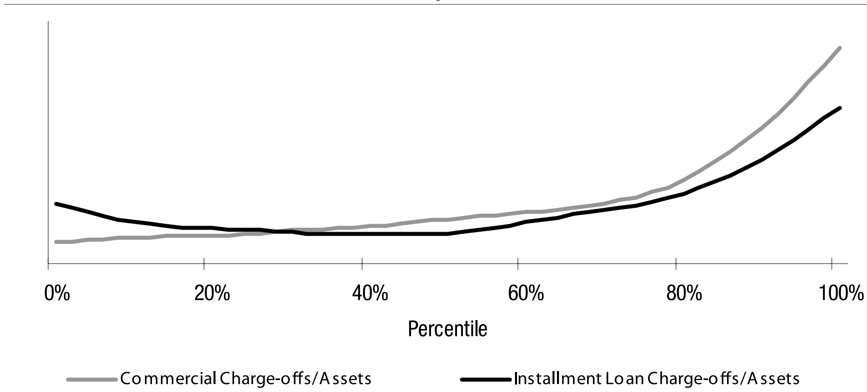
Commercial or business loans frequently comprise one of the most important assets of a bank. They may be secured or unsecured and for short or long-term maturities. Such loans include working capital advances, term loans, and loans to individuals for business purposes. As shown in **Figure 5**, the ratio C&I loans/Assets proved to be the most powerful predictor of bank default.

FIGURE 5: 5-Year Default Probabilities For More Concentration Ratios




Bank asset quality is commonly measured using the ratio of Loan Loss Reserve/Assets. The univariate relation between loan loss reserve ratio and default is indeed powerful but the ratio is highly correlated with other model variables (for example, a 33% correlation with C&I ratio). Additionally, we are more interested in measuring the loss flow due to non-performing loans rather than the historic asset performance. We found that two charge-off ratios (Commercial Charge-Offs/Assets and Installment Charge-Offs/Assets) have a powerful univariate relation with default and are not highly correlated with the other variables in the model.

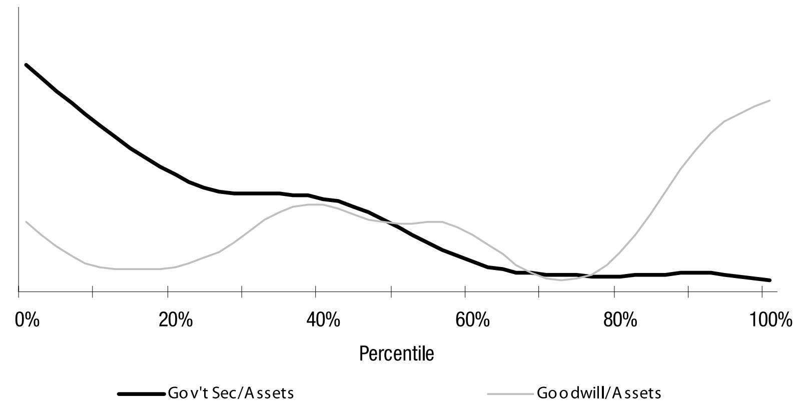
FIGURE 6: 5-Year Default Probabilities For Asset Quality Ratios




### 2.2.4 Liquidity

Liquidity risk refers to a bank's potential difficulties in meeting cash demands from current assets. The liquidity cushion – or technical reserve – held by a bank in addition to the regulatory requirement is maintained by most financial institutions in the form of treasuries, hence we measure liquidity as government securities/assets. The market for U.S. treasury securities is by far the largest, most active debt market in the world. Because of the extensive trading and degree of competition, dealers typically trade these securities at narrow bid-asked spreads. Government securities provide banks with liquidity at a relatively low cost. For example, a commonly used instrument for this purpose is the repurchase agreement in which the interest rate follows the federal funds rate closely. Financial institutions use this market as a source of overnight financing. Treasuries are also used as hedging instruments, to offset the interest rate risk inherent in posi- tions in other fixed-income securities. To incorporate a measure of liquidity risk, we use the ratio of Government Securities /Total Assets. **Figure 7** shows the relationship of this measure to default.

FIGURE 7: 5-Year Default Probabilities For Liquidity Ratios



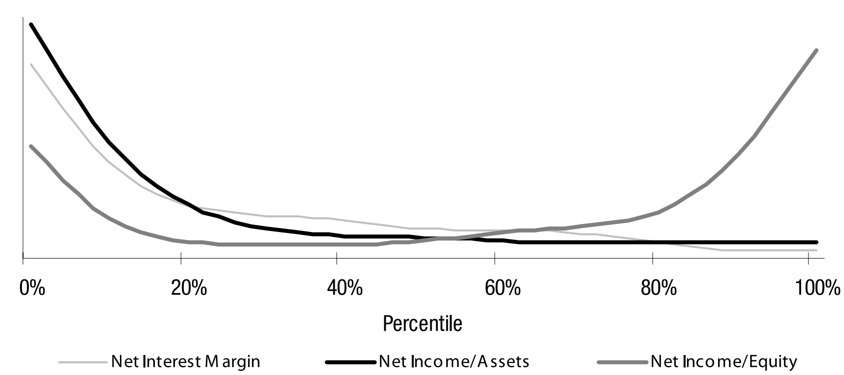

### 2.2.5 Profitability

As in many other industries, a broadly used measure of profitability in banking industry is the Return on Assets (ROA). In agreement with most of the empirical literature in this field, we found the univariate relationship between probability of default and ROA to be strong. Banks with higher earnings are, presumably, less likely to default. Clearly if the bank wishes to self-finance its growth (and in some cases to remain in the market) it needs capital. This additional capital can be obtained by retaining earnings or through investors that, in turn, require a minimum return to participate in the business. Since we are interested in business-related returns we deduct the extraordinary items from the numerator.

Another key factor, which affects profitability of a bank, is losses due to operational risk. Analysts often view the net interest margin as a measure of profitability and operational efficiency in the intermediation process. In this sense, a lower net interest margin implies a higher riskiness for the bank. Our model attempts to address the various types of risks a bank faces (credit risk, interest-rate risk, concentration risk, liquidity risk, and operating risk), for we acknowledge that a bank can fail from any particular one of these risks or from a combination of them.

**Figure 8** shows the relationship between several profitability measures of which Net Income/Total Assets shows the strongest relationship which is closely followed by the net interest margin.

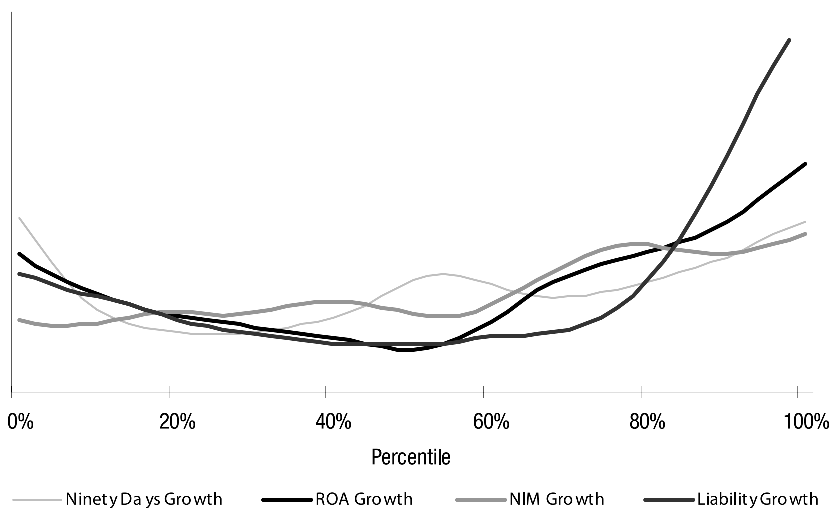
FIGURE 8: 5-Year Default Probabilities For Profitability Ratios



### 2.2.6 Growth

In general, empirical evidence has suggested that some banks whose growth rates were relatively high have experienced problems because their management and/or structure was not able to deal with and sustain exceptional growth. Banking is a highly leveraged business, and growth in liabilities provides good insight and is the most powerful in terms of probability of default relative to other measures of growth. As **Figure 9** shows, liability growth is somewhat u-shaped, as expected. We note that the high extreme (fast growth) is worse than the low extreme (slow growth) with respect to default probabilities. Moreover, we note that liability growth is a better default predictor relative to other candidate variables such as ROA and NIM growth, which exhibit a flatter u-shaped pattern in probability space.

FIGURE 9: 5-Year Default Probabilities For Growth Ratios




### 2.2.7 Macro Factors

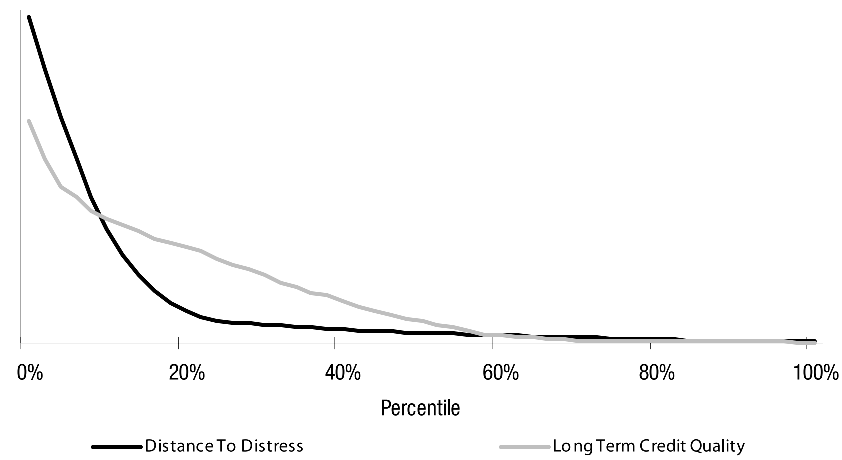
As mentioned earlier, there are differences between the causes of defaults in the banking industry and the corporate universe, where the former is primarily caused by financial (credit) risk, whereas the latter may be caused by business risk or financial risk. As a result, there is a notable distinction in the empirical default behavior of industrial firms and banks. Namely, default behavior moves in slow cycles for industrial firms while the same behavior for the banking industry is characterized by periods of low defaults that are followed by time periods with high clusters of defaults.

Therefore, from a modeling perspective, it is crucial for a default prediction model to be able to identify different economic environments. In order to capture this information, in addition to the company-specific factors, we include a macroeconomic variable in our model, to reflect information on changes in credit quality in the financial markets. Some potential time-series indicators to that end that come to mind are the *MBBI Index* (Moody's Bankrupt Bond Index), constant maturity Treasury Yield series behavior (GS1, GS2 and GS5), and Moody's *Trailing Speculative Default Rate Index*. Upon examination, we find that empirically all three measures are statistically significant and possess predictive power. Nevertheless, the MBBI index is a price index by construction, and thus, in addition to default rate information also incorporates information on recovery. Similarly, the constant maturity treasury yield series is responsive not only to the credit quality in the market, but also to other macroeconomic factors. Therefore, we chose to use the Moody's Trailing Speculative Default Rate Index in our estimations.

### 2.2.8 Distance to Default Measure And Long-Term Credit Quality Estimate

As mentioned earlier, there are a non-negligible number of privately-held banks that are actually subsidiaries of publicly-traded parent banks. The behavior of a subsidiary bank is often affected by the stock performance of its parent bank. In order to address this issue, we found it valuable to include the parent's equity market information in the default prediction model for banks with a publicly traded parent. We do this through the use of a contingent claims model of default that is motivated by Merton (1974). In addition, we include parent's long-term credit quality estimate as input for these banks.

FIGURE 10: 5-Year Default Probabilities For Distance To Distress And Credit Quality

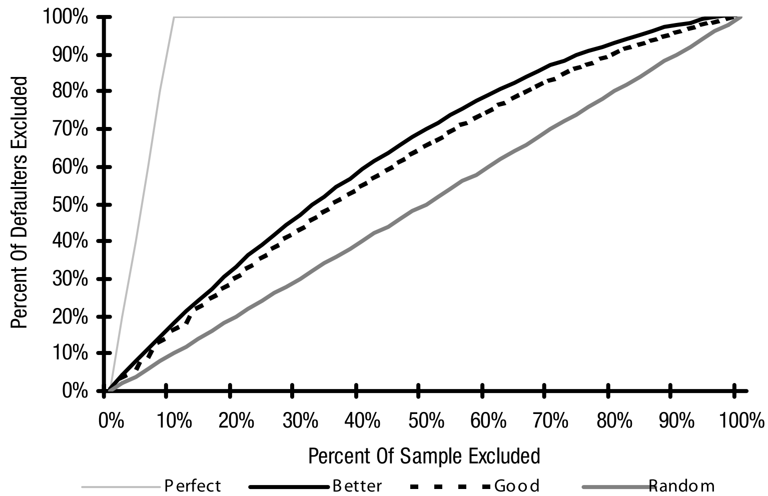



## 3. Empirical Results

As with previous RiskCalc models, the empirical tools we use for assessing statistical power, the ability to rank-order defaulters and nondefaulters, are power curves and accuracy ratios.

The intuition of a power curve can be described as follows: it maps the fraction of all companies with the worst score (horizontal axis) onto the fraction of defaulting companies within that group (vertical axis). Ideally, if the sample contained, for instance, 10% defaulters, then a perfect model would exclude all those defaulters at 10% of the sample excluded, where the lowest ranked companies would include all (and only) the defaults. Correspondingly, an uninformative random model would only exclude all of the defaulters by excluding the entire portfolio (see Figure 11).

FIGURE 11: An Illustration Of The Power Curve: Power




In reality, defaulters will not be perfectly discriminated, thus yielding a concave function: at 10% of the sample excluded 30% of the defaulters might be excluded, at 20% of the sample 40% of the defaulters might be excluded, and so forth.

It is important that models *predict*, not simply explain defaults, and therefore important to test the models using data received prior to the date of default. The average lifetime of a new loan is more than a year, so using financial statements several years prior to the default date is extremely useful.[11]

Based on power curves, we calculate the accuracy ratio (AR) for the estimated models. AR is simply the ratio of the area under the power curve of the estimated model to the area under the "ideal" model where the above random model has been subtracted from both. Accordingly, AR takes a value between 0 and 100%, where an AR of 0% indicates that the estimated model offers no improvement over a random model and an AR near 100% suggests that the estimated model is very close to the ideal model.

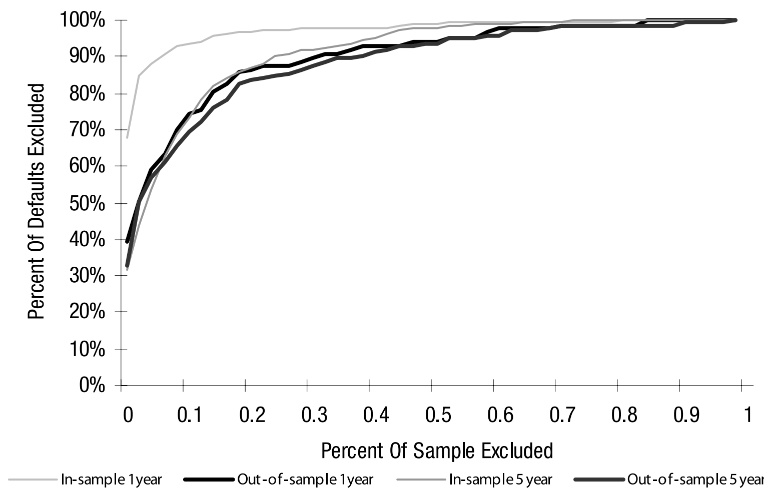
We present the results of one version of these results in **Figure 12**, which shows the in-sample performance of the model. There is evidence that the model exhibits high degrees of power in distinguishing good credits from bad ones. However, these are *in-sample* results. An immediate question one may raise is whether these performance statistics would hold for different segments of the overall sample. Put differently, we would like to know whether these results are robust throughout the sample and are not the result of overfitting the data. The best way to address these issues is to see how the model would have performed in the past against future data because we can compare its predictions against what actually happened. This is a walk-forward test.

In the walk-forward test, one estimates the model on the data up to a certain point in the past and scores the future year (relative to that point) with the model estimated. Then this point is advanced a year and the process is repeated. The process is continued in this manner until there is no future data available. Then all the scored out-of-sample subsamples are put together and accuracy ratio (AR) and power curves on the combined set is calculated. Note that this test is always out-of-sample for the data being tested.

The short- and long-run walk-forward power curves are displayed in **Figure 12**:

11This creates a complication because if we simply use each financial statement as an observation, we will double and perhaps triple-count defaulting firms: once for each statement prior to the default date. Our method of testing accommodates this complication by going backward in time from the default date, as opposed to forward in time from each statement date. This ensures that each failed firm is counted only once and does not bias our results. The details of the procedure by which we estimate both 1-year and 5-year default probability curves are described in the Appendix.

FIGURE 12: Power Of The RiskCalc Banking Model




As the graph illustrates, the model has high predictive power. The corresponding accuracy ratios for the one- and five-year models based on the walk-forward tests are 78.0% and 76.2%, respectively. These high accuracy ratios (ARs) are not surprising as the corresponding in-sample ARs are about 90.9% for 1-year horizon and 78.8% for 5-year horizon, respectively.

Another standard robustness test is "*k*-fold test." In order to implement this test, we divided the failing and non-failing banks into *k*-equally-sized segments. This yielded *k* equally-sized observation subsamples that exhibit the identical overall default rate and are temporally and cross-sectionally independent. Accordingly, we estimate the model on *k-1* sub-samples, and score the *k-th* subsample. We repeat this procedure for all possible combinations, and put the *k* scored "out-of-sample" subsamples together and calculate an accuracy ratio (AR) on this combined data set. Note that this test is always out of sample for the data being tested.

The table below summarizes the *k*-fold test results. The reported figures are the Accuracy Ratios (AR) by corresponding category and time spans.

TABLE 2: K-Fold Test Results

| 3-fold | | 5-fold | | 10-fold | |
|---|---|---|---|---|---|
| 1 year AR | 5 year AR | 1 year AR | 5 year AR | 1 year AR | 5 year AR |
| 88.1% | 60.3% | 88.2% | 61.2% | 87.4% | 61.4% |

Note that the high accuracy ratios of the RiskCalc Model for the U.S. Banks should not be interpreted as a "better" model in relation to the previous RiskCalc models, as model power is strictly a data-driven measure and thus, cannot be compared when ARs are estimated on different datasets and universes.

Since banks exhibit a wide variety in scope, type, and size, examining the predictive power of the model performs an additional check on the model by the charter class. **Table 3** below summarizes the results on the accuracy ratio of the model by bank charter class. As the table illustrates the model power seems to be very robust across different classes, and the 1-year model power is equal to or above 90% in each case. Similarly, 5-year model power is also relatively high for each subsample. Not surprisingly, commercial banks segmentation reveals a higher accuracy ratio in comparison to savings banks and associations. Nevertheless, recalling that the model was estimated on the commercial bank universe (in order to avoid any biases, which may be induced by the thrift clash in the 1980s), these statistics should be interpreted to reveal purely "out-of-sample" performance of the model on these types of institutions.

TABLE 3: Performance Of RiskCalc Model By Bank Charter Class

| 1 Year Power | 5 Year Power | Code | Bank Charter |
|---|---|---|---|
| 91% | 76% | N | Commercial bank, national (federal) charter and Fed member, supervised by the Office of the Comptroller of the Currency (OCC) |
| 93% | 84% | SM | Commercial bank, state charter and Fed member, supervised by the Federal Reserve (FRB) |
| 90% | 79% | NM | Commercial bank, state charter and Fed nonmember, supervised by the FDIC |
| 94% | 65% | SB | Savings banks, state charter, supervised by the FDIC |
| 90% | 64% | SA | Savings associations, state or federal charter, supervised by the Office of Thrift Supervision (OTS) |

In order to test the out-of-sample performance of the RiskCalc model by charter type, we also conducted a walk-forward test on each of the categories above (except savings associations since the AR figure above is already out of sample for that universe by design). The corresponding results are outlined in Table 4 below:

TABLE 4: Walk-Forward Power Tests By Bank Charter

| 1 Year AR | 5 Year AR | Bank Charter Code |
|---|---|---|
| 85% | 73% | N |
| 88% | 77% | SM |
| 83% | 71% | NM |
| 89% | 80% | SB |

### Comparison With A Benchmark Model

In order to better assess the predictive power of the RiskCalc model, we implemented a recent empirical bank default model by Estrella et al. (2000) on the same data set and subjected it to the same set of out-of- sample and out-of-time walk-forward tests. Hence, since Estrella et al. and RiskCalc models are estimated on the same datasets, the estimated accuracy ratios can be compared with each other and thus the Estrella model performance can give some insight into the relative power of RiskCalc. It is worth pointing out at the outset that Estrella et al (2000) uses financial ratios without any transformation, and that the model coefficients vary in magnitude and/or sign from an estimation period to another by a wide margin. The corresponding one- and five-year ARs for the Estrella model turn out to be 63.8% and 35.7% (as opposed to 78% and 76% of RiskCalc walk-forward test results), respectively, which reaffirms the superiority of the RiskCalc Banking model in ranking ability and robustness over time.

TABLE 5: Estrella Et Al (2000) Versus RiskCalc Model: Out-Of Sample Performance

| Model | Estrella et al (2000) Model Power | RiskCalc - Walk-Forward Power |
|---|---|---|
| 1-year AR | 63.8% | 78% |
| 5-year AR | 35.7% | 76% |

### 3.1 MODEL WEIGHTS

In addition to studying the technical methodology of the model, appropriate use of any quantitative model calls for a solid understanding of the intuition behind it. In other words, in order to use and interpret the model output, the analyst needs to know which inputs drive the model, and their respective weights in determining the model's score. This knowledge is important for understanding which factors may be contributing to the level of a PD calculated for a given firm and evaluate the sensitivity of the score.

In particular, the relative weights for each input are calculated as follows. First, the model inputs that are transformed into the probability space are shocked for one standard deviation. The corresponding PDs are calculated for each individual variable and expressed as deviations from the mean. The weight of each input is obtained by dividing these individual deviations by the summation of the absolute deviations of all of the model inputs. By definition, these weights add up to 100%.

The calculated weights of input variables for the one- and five-year models are presented in **Table 6** by category.[12]

TABLE 6: Model Weights

| Category | One-Year Model | Five-Year Model |
|---|---|---|
| Concentration | 31% | 40% |
| Profitability | 27% | 22% |
| Capital | 22% | 13% |
| Liquidity | 7% | 8% |
| Asset Quality | 5% | 11% |
| Overall Credit Quality | 8% | 4% |
| Liability Growth | | 2% |

As **Table 6** illustrates, the loan portfolio structure seems to be an important factor in determining bank defaults: combination of different types of loans (e.g., C&I loans, real estate loans, and construction loans make up to 40% (31%) of the total weight in the 5-year (1-year) model). As in many other default prediction models, profitability turns out to be a significant component (22% in the long term model and 27% in the short term model), and capital structure of the banks is also seen to be a key component (13% in 5-year model, 22% in 1-year model). Risk-free and liquid assets proxied by the portion of government securities expressed as a percentage of assets also is significant in both the short and long term models (7% in the short term and 8% in the long term model).

Net interest margin (NIM) seems to be more influential in longer horizons rather than the short term; its weights are 6% in the 5-year model and 2% in the short term model. Charge-offs also provide a statistically significant signal, but as in the case of NIM, their effect is more pronounced in the long term: 11% in the long term and 5% in the short term. Liability growth is observed to have about 2% influence in the 5-year model. Finally, a change in the overall credit quality in the economy, as measured by the deviation of the 24-month trailing default rate index, is seen to be significant in both models; not surprisingly its relative weight is higher in the short term (8%) vis-à-vis the long term (4%).

12 See Table 1 for the model variables in each firm-specific ratio category. Overall Credit Quality is proxied by the trailing default rate variable.

For banks with a publicly traded parent we also have distance to default and long-term credit quality estimate as inputs. In this light, it is interesting to note that the weights calculated reflect the relative importance of equity information and the fundamentals. In the short run, the distance to default measure has a weight of 51%, whereas the financial ratios carry a weight of 42%. In the long run, however, the roles seem to be reversed so that in the 5-year model the financial ratios accounts for 47% of the variation, where the distance to default has a weight of only about 41%. Long-term credit quality estimates account for 12% of the variation in the 5-year model and about 7% of the variation in the 1-year model.

## 4. Data Used In The Development Of RiskCalc For U.S. Banks

The RiskCalc U.S. Bank model was constructed using information on 17,673 unique financial institutions for the 1986-1999 time period. A total of 161,034 observations was used.

We use the FDIC definition of default in our modeling exercise, and similar to Wheelock and Wilson (2000), also include FDIC financial assistance as a default event. In particular, the FDIC determines three types of transactions as defaults.

"Failing institutions have been resolved through several different types of transactions. The transaction types outlined below can be grouped into three general categories, based upon the method employed to protect insured depositors and how each transaction affects a failed institution's charter.

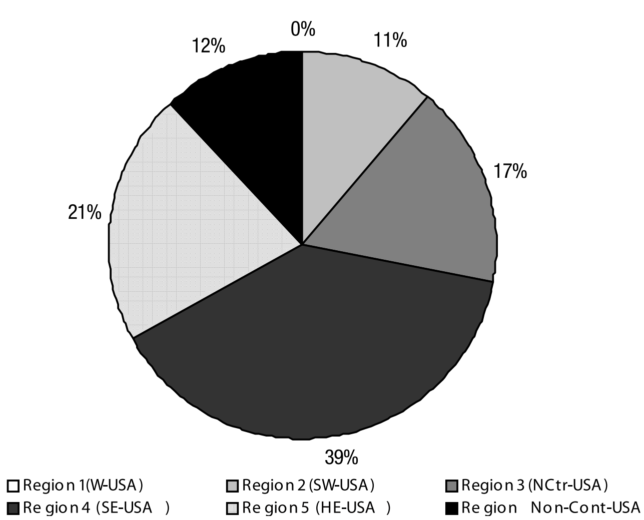
» **Assistance Transactions.** In most assistance transactions, insured and uninsured depositors are protected, the failed institution remains open and its charter survives the resolution process.

» **Purchase And Assumption Transactions.** In purchase and assumption transactions, the failed institution's insured deposits are transferred to a successor institution, and its charter is closed. In most of these transactions, additional liabilities and assets are also transferred to the successor institution.

» **Payoff Transactions.** In payoff transactions, the deposit insurer - the FDIC or the former Federal Savings and Loan Insurance Corporation - pays insured depositors, the failed institution's charter is closed, and there is no successor institution."

**Figure 13** illustrates the regional distribution of banks in the data set. The figure shows that majority of banks are concentrated in the Northcentral (Region 3), Southeastern USA (Region 4) and Southwestern USA (Region 2)[13].

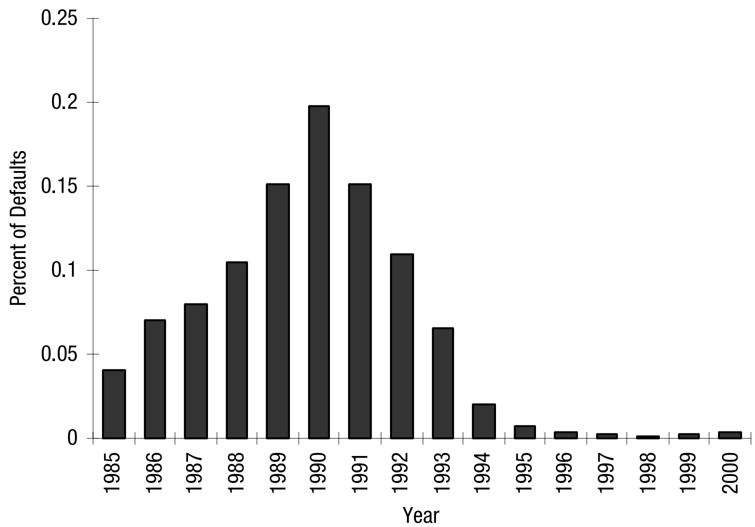
FIGURE 13: Regional Bank Distribution




13 For a complete definition of regions by states please refer to Appendix D.

As of June 2000, our database contained more than 2,400 (public and private) defaults in the banking industry; where bank defaults are characterized according to FDIC descriptions and include assisted mergers and payoffs.

FIGURE 14: Bank Defaults




As shown in **Figure 14**, the bulk of defaults in our dataset (49%) occurred during the 1989-1991 period. We note a disproportionate number of thrift defaults (46.2%).[14] The majority of the failed institutions were Savings Associations, supervised by the Office of Thrift Supervision, followed by the State Banks, which are not members of the Federal Reserve system and are supervised by the FDIC. National Commercial banks, which are supervised by the OCC, made up the third largest group in the list of defaults for the 1984-1999 period. Over one third of those defaults occurred in the South Central Region of the USA (Region 2), where Texas accounts for 26% of the total defaults – followed by the Southwest Region (Region 4) –mainly Lousiana and Florida – and North Central Region (Region 3). The Northeast (Region 6) was the least affected. As observed by the FDIC[15] the bank defaults were concentrated in relatively few regions. Of the total defaults, nearly 50% were concentrated in five states: Texas, California, Louisiana, Oklahoma and Florida. Included in these are defaults on banking holding companies - not subsidiaries.

The FDIC summarizes the events associated with bank defaults in these particular states as:

» severe economic downturn related to the collapse in energy prices (Louisiana);

» agricultural recession of the early 1980s (Oklahoma and Texas);

» an influx of banks chartered in the 1980s (California and Texas).

Of particular interest is Texas, which was affected simultaneously by several of these factors: The rapid rise and subsequent collapse of oil prices, the commercial real state boom and bust, the effects of the agricultural recession, and state prohibitions against branching are, according to the FDIC, key points in explaining the high volume of financial institutions that defaulted during the period of this study.

In order to avoid any systematic bias in our modeling efforts we excluded thrift defaults from our model estimation sample and subsequently tested the performance on each banking charter segment (e.g. commercial banks that are supervised by OCC, commercial banks that are supervised by FRB, commercial

14 For the scope of this model we focus on banks with an asset size of equal or in excess to 100 million dollars.
15 History of the Eighties-Lessons for the Future; Chapter 1

banks that are supervised by FDIC, savings banks that are supervised by the FDIC, and savings associations that are supervised by the OTS) to ascertain that the model performance is robust throughout the various market segments. Thus although we did not fit the models using thrifts, we did include them in our tests of the models.

After correcting our dataset for double-counting due to defaults of the same institution in several locations we were able to use 413 private bank defaults suitable for modeling. As **Figure 14** shows, the resulting defaults are fairly evenly distributed in time and capture the "default wave" of late 1980s and early 1990s. Moreover, since the data also covers the post-default wave period (i.e., late 1990s and 2000), the model coverage is unbiased and reflects an entire cycle in terms of historical data.

## 5. Potential Uses Of RiskCalc Model And Target Population

As with other RiskCalc models, RiskCalc for U.S. Banks provides a fast and efficient credit quality scoring tool, which can be implemented on both individual names as well as large portfolios of credits. Thus, by its very nature it can be used for a wide variety of uses ranging from individual decision making to portfolio risk monitoring, loan pricing, as well as CDO analysis.

Default probability estimates, such as those produced by RiskCalc, are at the heart of the commercial credit evaluation process. In financial institutions, the margin between incoming and outgoing cash flow can be thin and the leverage high, so that small differences in asset quality affect their solvency, and thus the solvency of financial systems. Benchmarks, such as those generated by RiskCalc, provide a way for market participants to evaluate the credit quality of different financial institutions more effectively.

Regulatory agencies categorize and treat banks differently depending upon their asset size. Banks with assets less than $100 million file call report form FFIEC 034, while larger banks file consolidated form FFIEC 033, FFIEC 032 if they have domestic offices only, or FFIEC 031 if they have domestic and foreign offices.[16] From a modeling perspective, the regulatory difference in treatment can perhaps be justified on the empirically observed differences in the financial ratios in the universe of banks with less than $100 million in assets and those above this limit: we observe that banks with assets in excess of $100 millions tend to have notably higher treasury holdings relative to their assets, higher returns on assets, and higher NIM ratios. Moreover, they also have higher commercial and industrial and real estate loans relative to their asset holdings. Thus, in our model we went along with that distinction and set $100 million in asset value as the lower threshold for applicability.

The RiskCalc model for privately-held U.S. banks is intended to cover commercial banks, thrifts as well as holding companies. The target universe for this model is institutions with assets above $100 million that have no publicly traded equity at the bank level. Thus, the model is not optimally suited for a firm listed on a stock exchange or even one with shares traded over-the-counter, because the model ignores useful predictive information (i.e., market value and equity price volatility). Publicly traded banks are more meaningfully scored by Moody's KMV's Financial Firm Model.

### 5.1 SOME IMPLEMENTATION RECOMMENDATIONS

There are some conceptual issues to consider when implementing RiskCalc for U.S. Banks. First, RiskCalc was designed with a bias towards parsimony. Our rule of thumb is to keep the model size small and not to include a variable unless we can statistically demonstrate its predictive power in the univariate and more importantly in the multivariate context and it resulted in improved predictive power. The model also tends to be empirically conservative in variable selection and does not include variables for which we do not have full historical data coverage.

16 Similarly, Bank Holding Company report forms: Y-9C , Y-9LP, Y-9SP, and thrifts file TFR forms.

For example, the model does not include some financial ratios that are based on BIS regulatory disclosure information (e.g., capital adequacy, which utilize data that has been available only for a limited number of years). In future versions, Moody's KMV clearly will increasingly refine the model to account for such factors where appropriate. Our aim in developing the RiskCalc network of models is not merely to provide a set of powerful tools but also to ensure that they can be used without imposing overly burdensome data requirements on users.

Another issue is the nature of analysts' versus RiskCalc scores in general. As with any other quantitative model, RiskCalc produces PD estimates given a subset of publicly available quantitative information in the market place. To the extent that this information reflects accurate assessments of the companies' managerial quality, business plans, options, and so forth, RiskCalc PD estimates will not notably diverge from analysts' views on them. If one has additional information, we recommend its incorporation it into the analysis and modification of the RiskCalc PDs. For example, one can conduct a scenario analysis by changing the values of the affected variables and keeping everything else constant and observing the corresponding PD estimates.

## 6. Conclusion

The RiskCalc methodology is based on sound theory and practical modeling and credit experience. The model is econometrically derived, well understood, and sophisticatedly simple, relying on well-established risk factors. By transforming, or "mini-modeling," the input ratios and then combining them into a multivariate model, we capture and integrate a non-linear problem, yet retain transparency. The final mapping process takes into account our central tendency view of default rates.

We see default modeling as a forward-looking problem, and thus apply a rigorous validation methodology to check for robustness. The final result is a model that we believe is well suited to forecast future defaults, rather than only statistically explain previous ones.

Using RiskCalc for U.S. Banks should help improve profitability through the credit cycle, from credit decisioning to pricing to monitoring to securitizing. Clearly, RiskCalc is not intended as a sole measure of overall risk; nevertheless, it should be viewed as a very powerful aggregator of financial statement information into a meaningful and validated number that allows consistent comparison of portfolio risks.

## 7. Appendix A: Power Curves

A power curve[17] is constructed by plotting, for each threshold, the proportion of defaults excluded at various levels of sample exclusion. The vertical axis measures the percent of defaults excluded conditional upon excluding various percentage levels of the sample. Thus if using a score to exclude 50% of the sample caused it to lose 80% of the defaulting companies, the power curve would go through a line corresponding to x=0.5 and y=0.8.

$$power\ (b) = \frac{\sum_{t=1}^{b} p(t)}{\sum_{t=1}^{B} p(t)} = \frac{defaults\ excluded\ at\ b}{total\ defaults} \qquad (A.1)$$

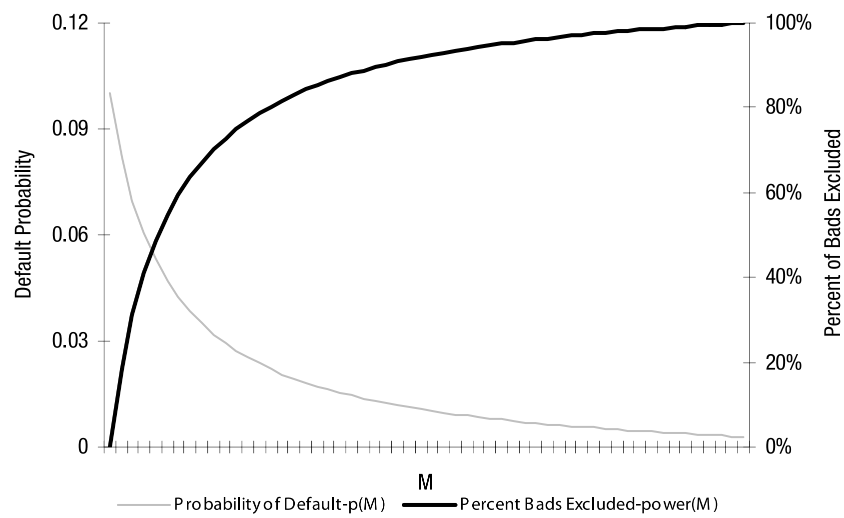
Here B is the total number of bins (often 10 for illustrative purposes), and b is a particular bin. The power at bin b represents the sum of all the defaults in the 'worst' fraction b/B of the scores, as ranked by the metric M.[18]

17 Versions of this measure are also known as Gini curve, CAP plot, Lorenz curve, ordinal dominance graph, or ROC curve.

18 Since defaults excluded 'at bin b' is ambiguous – it could mean 'up to bin b' or 'up to and including bin b' – we calculate the area using the average of the two methods. Nevertheless this adjustment makes practically no difference.

The net result is **Figure A.1** below, which shows the probability of default for a level of M, and statistical power, which pertains to the nature of the data up to a level of M. In this case we rank-order the firms from risky (left) to less risky (right), so that the P(M) and Power(M) correspond. The graph shows a particular case. This type of model would quickly have excluded most of the bad companies: a 20% exclusion of the worst companies according to the M score would exclude 70% of the future defaulters.

FIGURE A.1: Power Curve And PD Curve




There is a one-to-one correspondence of power and probability of default by rank order, in that for any point t along a default metric:

$$p(t) = \bar{p} * \frac{\partial\, power\ (t)}{\partial t}, \qquad (A.2)$$

where $\bar{p}$ is the mean probability of default.

While the graphical or tabular display of power is informative and has the advantage of allowing one to examine power at a variety of thresholds, it is useful to aggregate the power curve information into a single number that allow unambiguous comparison. One such measure is the area under the power curve. A model more bowed out towards the left will have a greater area, and be more powerful on average. Using the area under the power curve implies that there can exist threshold levels such that a model with a smaller total area has a momentary advantage. Thus the area is not a measure of global or complete dominance, just an intuitive measure of dominance on average. The area can be calculated using equation (A.1) above, specifically:

$$Area = \frac{1}{B}\sum_{b=1}^{B} power(b) \qquad (A.3)$$

where B is the total number of bins. If the Area is greater for one model than another, it is more powerful.[19]

19 Clearly in order to compare models one not only needs an aggregate measure of power, but also a standard error on this metric. For further details on model validation and error bounds please see: Benchmarking Default Prediction Models: Pitfalls and Remedies in Model Validation (Stein, 2002)

## 8. Appendix B: Calibration Curve Construction Details

The model was calibrated to a one-year and a cumulative five-year horizon. For the one-year horizon we took each defaulting firm, and found its score as of last available statement prior to the default date. The motivation for moving back in time was not to score institutions which have already defaulted.

If no score is available 3-27 months prior to default date, we exclude the observation. Where a score is available, we map it to a percentile, and this collection of percentiles forms the basis from which the power curve used in calibration is created.

Specifically, given a collection of percentiles of defaulting firms $\{\phi_j\}^{J}_{=1}$, where J is the total number of defaulting firms, the power for each bin *(b)* is simply:

$$power(b) = \frac{1}{J}\sum_{j=1}^{J}\left\{1 \mid \phi_j \leq \frac{b}{B}\right\},$$

where B is the total number of bins, and $\{1 \mid \phi_j \leq \frac{b}{B}\}$ is an indicator function equal to 1 if the defaulting firm, *j*, was in a percentile lower than *b/B*.

For example, for a one-year power curve, we would take a default in 7/98, and move back to 4/98 to find the percentile of the RiskCalc score using that month. As is most probable, the statement date is not exactly at 4/98, and so we must go back in time, to 3/98, then 2/98, etc., until we find the date at which we have a financial statement.

The five-year default curve was constructed using the average power curves of five cohorts starting in 1993, 1994, 1995, 1996 and 1997. A score within a given cohort was flagged as defaulting if it defaulted at some point between 3 and 63 months after the date of the financial statement. We did not flag companies that defaulted within the first 3 months following the date of the financial statement date. This was done to take into account the lag between the publication and the date of the financial statement, and in order to reflect that we use a relatively late definition of default.[20]

## 9. Appendix C: Previous Bank Default Modeling Literature

Banking is the most highly regulated industry in the world because, as pointed out by Gilbert, Meyer and Vaughan (2000), bank defaults have stronger adverse effects on economic activity than other business default. Throughout history, bank defaults have caused serious disturbances in monetary and real markets. Bank defaults hinder the flow of credit, disrupt the payment chain, and reduce money supply. Several authors (e.g., Friedman and Schwartz (1971)) believe the great depression in the United States in the 1930s was exacerbated by the disruption of credit caused by bank defaults.

Similarly, in international markets, bank defaults have caused notable economic disturbance. The default of Bankhaus Herstatt in West Germany at the end of 1974, for example, triggered turbulence in international currency and banking markets which contributed to the creation of the Basel Committee.

During the 1980s, the U.S. banking system experienced a significant increase in the number of banking defaults. As observed by Wheelock and Wilson (2000), since 1984 “the number of commercial banks in the United States has fallen by one-third, reflecting first a wave of defaults and then, more recently, unprecedented number of acquisitions and mergers.” In order to allocate scarce resources more efficiently, the U.S. regulatory agencies increasingly incorporate different econometric models in their surveillance practices, improving their early warning system. Hence, quantitative models have now become an important complement to the traditional supervisory screens and help the regulators to narrow the scope in the detection of financially weak institutions.

20 For details on standard error of AR calculations please see: Stein (2002).

In the construction of some of the early models, Bovenzi, Marino and McFadden (1983) chose a probit framework to estimate the PD for U.S. commercial banks. The selected variables account for the inherent risks in the commercial bank line of business (i.e., credit risk, portfolio diversification, internal controls, operative inefficiencies, capitalization, and interest rates). In their definition of "default" the authors include banks that receive funds from the Federal Deposit Insurance Corporation (FDIC). Two sources of information were used to build the models. The first was information sent by the Banks to the supervisory office and the second was additional information received from the *in situ* inspection. Those independent sources of information were used to estimate PD for one-, two- and three-year horizons. For the one- and two-year horizon, the model that includes additional information has more predictive power than the model that includes standard information, but not for the three-year horizon.

Gilbert, Meyer and Vaughan (1999) developed a model that combines financial ratios to estimate the probability of default within the next two years for Fed-supervised banks. In their work, the authors used 1987, 1988 and 1989 call report data to predict banking defaults during 1989, 1990 and 1991 respectively, and estimated a logit model. The authors identified a set of financial ratios – focusing on capital adequacy, asset quality, managerial competence, earnings strength, liquidity risk and control variables – that were used commonly by officers and examiners in the eight Federal Reserve Districts and in the reviewed literature. Fourteen financial ratios[21] were defined as relevant in predicting PD. They concluded that the use of these models improved the surveillance practices when used as complements with the traditional supervisory screens by predicting changes in the bank's CAMEL rating. At the same time, they showed that the hypothesis that their econometric model has no explanatory value was "soundly rejected for all seven years."

Similarly, Estrella, Park, and Peristiani (2000) developed a bank default prediction model that utilized similar ratio categories. Based on their findings, they concluded that relatively "simple" ratios (e.g., core capital over gross revenues, total tangible assets), performed very well over a one- to two-year horizon). Only for predictions over longer periods do more "complex" ratios (e.g., core capital over total risk based capital, perform relatively better).

Wheelock and Wilson (2000) choose the competing-risk hazard models framework to estimate both the probability of being acquired and of failing. The authors use a database maintained by the Federal Reserve and provided by the National Information Center (NIC) that include quarterly statements and call reports filed by commercial banks. Data on 3,972 banks and 230 bank defaults was used in the hazard-estimation model. Their sample accounts for all banks in existence in 1984 with at least $50 million of assets. Their empirical model was based on measures used by regulators to evaluate a bank: capital adequacy, asset quality, management, earnings and liquidity. They found that banks with high leverage, low earnings, low liquidity, or risky asset portfolios were more likely to fail. Geographic diversification was also found to be an important characteristic, as they conclude that banks located in states that allow interstate branching were also less likely to fail. In their effort to model management quality they use three inefficiency measures: (i) cost efficiency, (ii) input distance function measure of technical inefficiency, and (iii) inverse of output distance function measure of technical inefficiency. A parametric stochastic frontier was defined to estimate measure one and for the remainders non-parametric distance functions were utilized.

21 The ratios are: total equity to total assets, non-performing loans to total loans, consumer loans to total assets, other real estate owned to total loans, non interest expense to total revenue, insider loans to total assets, occupancy expense to average assets, return on assets, interest income accrued to total loans, liquid assets to total assets, large time deposits to total assets, core deposits to total assets, natural logarithm of total assets, and total assets to total assets in the parent-holding company.

## 10. Appendix D: Regional Mapping of States

| Region | State Name |
|---|---|
| 1 | California |
| 1 | Colorado |
| 1 | Idaho |
| 1 | Montana |
| 1 | Nevada |
| 1 | Oregon |
| 1 | Utah |
| 1 | Washington |
| 1 | Wyoming |
| 2 | Arizona |
| 2 | New Mexico |
| 2 | Oklahoma |
| 2 | Texas |
| 3 | Illinois |
| 3 | Indiana |
| 3 | Iowa |
| 3 | Kansas |
| 3 | Michigan |
| 3 | Minnesota |
| 3 | Missouri |
| 3 | Nebraska |
| 3 | North Dakota |
| 3 | Ohio |
| 3 | South Dakota |
| 3 | Wisconsin |
| 4 | Alabama |
| 4 | Arkansas |
| 4 | Florida |
| 4 | Georgia |
| 4 | Kentucky |
| 4 | Louisiana |
| 4 | Maryland |
| 4 | Mississippi |
| 4 | North Carolina |
| 4 | South Carolina |
| 4 | Tennessee |
| 4 | Virginia |
| 5 | Connecticut |
| 5 | Delaware |
| 5 | District Of Columbia |

| | |
|---|---|
| 5 | Maine |
| 5 | Massachusetts |
| 5 | New Hampshire |
| 5 | New Jersey |
| 5 | New York |
| 5 | Pennsylvania |
| 5 | Rhode Island |
| 5 | Vermont |
| 5 | West Virginia |
| 5 | Puerto Rico |
| 6 | Alaska |
| 6 | Hawaii |

## References


1. Basel Committee on Banking Supervision, "A new Capital Adequacy Framework." Basel, June 1999.
2. Bovenzi, J. F., J. A. Marino, and F. E. McFadden. "Commercial bank failure prediction models, Economic Review." *Federal Reserve Bank of Atlanta* November 1983: 14-26.
3. Dupont, D., and B. Sack. "The Treasury Securities Market: Overview and Recent Developments." *Federal Reserve Bulletin* December 1999.
4. Estrella, Arturo, Sangkyun Park, and Stavros Peristiani. "Capital Ratios as Predictors of Bank Failure." *Federal Reserve Bank of New York Economic Policy Review* July 2000.
5. Falkenstein, E., A. Boral, A., and A. E. Kocagil. *RiskCalc for Private Companies II: More Results and the Australian Model*, Moody's Investors Service Special Comment, December 2000.
6. Falkenstein, E., L. V. Carty, and A. Boral. *RiskCalc for Private Companies: Moody's Default Model*, Moody's Investors Service Special Comment, May 2000.
7. FDIC, "History of the Eighties - Lessons for the Future Volume I: An Examination of the Banking Crises of the 1980s and Early 1990s." Division of Research and Statistics, 1-84.
8. Friedman, M., and A. J. Schwartz. *Monetary History of the United States*, 1867-1960. Princeton: Princeton UP, 1971
9. Gilbert, R. A., A. P. Meyer, and M. D. Vaughan. "The role of a Camel Downgrade Model in Bank Surveillance." *The Federal Reserve Bank of St. Louis* August 2000.
10. Gilbert, R. A., A. P. Meyer, and M. D. Vaughan. "The role of Supervisory Screens and Econometric Models in Off-site Surveillance." *Federal Reserve Bank of St. Louis* November/December 1999.
11. Meyer, Paul A., and Howard W. Pifer. "Prediction of Bank Failures." *Journal of Finance* September 1970: V25, number 4.
12. Merton, R. C. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates." *Journal of Finance* 29 (1974): 449-470.
13. Peek, Joe, and Eric S. Rosengren. "How Well-Capitalized are Well-Capitized Banks?" *New England Economic Review*, September-October 1997: 41-50.
14. Sobehart, J.R., S. C. Keenan, and R. Stein. *Benchmarking Quantitative Default Risk Models: A Validation Methodology*, Moody's Investors Service Special Comment, March 2000.
15. Stein, R. *Benchmarking Default Prediction Models: Pitfalls and Remedies in Model Validation*, Moody's KMV, Technical Report, 2002.
16. Thompson, James B. "Predicting Bank Failures in the 1980s." *Federal Reserve Bank of Cleveland Economic Review*. First Quarter 1991: 9-20.
17. Wheelock, David C., and Paul W. Wilson. "Why do Banks Disappear? The determinants of U.S. Bank Failures and Acquisitions." *The Review of Economics and Statistics* February 2000