# Bank Failure Prediction Using DEA to Measure Management Quality

Richard S. Barr* Thomas F. Siems†

June, 1996

**Abstract**

Presented are new failure-prediction models for detecting a bank's troubled status up to two years prior to insolvency using publicly available data and a new category of explanatory variable to capture the elusive, yet crucial, element of institutional success: management quality. Quality is assessed using data envelopment analysis (DEA), which views a bank as transforming multiple inputs into multiple outputs, focusing on the primary functions of attracting deposits and making loans. The new models are robust and accurate, as verified by in-depth empirical evaluations, cost sensitivity analyses, and comparisons with other published approaches.

Bank failures have increased at an alarming rate in the past decade. The number of commercial bank failures averaged less than seven per year from 1950 to 1980, but has since escalated to an annual average of 175 from 1986 through 1991. This rapid increase raises many questions regarding the safety and soundness of the banking industry: Why has there been a sudden increase in the number of bank failures? What can be done to slow the bank failure rate? How can the likely collapse of an institution be anticipated and prevented?

Often the failure of a bank can be avoided, or the bailout costs minimized, by early detection of a bank's troubled status and subsequent intervention by regulatory authorities. Key to this effort is the identification of the bank's potential for failure. To this end, researchers have sought mathematical models that predict institutional failure in an accurate and timely manner.

Interest in predicting bank failures dates from Secrist's classic study [21]. It examined 741 national banks that failed in the late 1920s and early 1930s and 111 banks that did not fail prior to 1933 to "secure indications of likely survival or of death." This comparative analysis was the first of its kind and sought to discover the symptoms of failure and non-failure.

*Department of Computer Science and Engineering, Southern Methodist University, Dallas, Texas 75275, email: barr@seas.smu.edu.

†Financial Industry Studies, Federal Reserve Bank of Dallas, Dallas, Texas 75201, email: siems@seas.smu.edu.

Not until the early 1970s did interest in predicting bank failures return. This renewed interest stemmed from the development of "newer" classification techniques, such as discriminant analysis and nonlinear regression, for separating "potential failures" from "probable survivors." Now, with the number of failures surging, the need for more effective prediction models has become evident. While an early-warning system cannot replace the on-site examination, which allows for personal interaction with the bank's management and employees and permits first-hand evaluation of operating procedures, levels of risk-taking, and long-range strategic planning. However, an effective warning system can complement the on-site examination process by identifying troubled institutions that need early examination or possible intervention.

As outlined in [25], for an early-warning (failure-prediction) model to be useful to regulators it must have the following ingredients:

- *Understandability.* For the model to be accepted and used by the examination staff, it must be understandable and fit their intuitive perceptions of how banks fail; the model will not be trusted if it is considered a "black box" where banks are rated according to some highly complex process.

- *Ease of use/quickness.* To yield the greatest benefits, an early-warning system's data must be timely and accurate . Fortunately, banks are already required to submit quarterly balance sheet and income statement information to banking regulators, and models built around this database can rely on up-to-date, reliable inputs.

One means of enhancing the regulators' understanding of a model is to build it around familiar factors. When examiners evaluate a bank's health, they develop an overall rating based on *C*apital adequacy, *A*sset quality, *M*anagement quality, *E*arnings ability, and *L*iquidity position—hence the term *CAMEL rating.* Federal regulators developed the numerical CAMEL rating system in the early 1970s to help structure their examination process. Since then, the use of CAMEL factors has become widespread, due to its simplicity and use by regulators. Financial data and relationships are the principal ingredients for scoring capital, asset quality, earnings, and liquidity. Assessing management quality, however, requires professional judgment of a bank's compliance to policies and procedures, aptitude for risk-taking, development of strategic plans, and the degree of involvement by the bank's officers and directors in making decisions.

Most failure-prediction models include variables that can be categorized under four of the five CAMEL factors; the variable that is typically missing is the one which assesses management quality. This is a rather curious paradox since the quality of management is often cited as the *leading* reason for failure.

In this paper, we present a bank failure prediction model that is developed around a new paradigm for assessing a bank's management quality. This paradigm views a bank as processing multiple inputs to produce multiple outputs, and focuses on its key financial intermediation functions of acquiring deposits and making loans and investments. Using *data envelopment analysis* (DEA), a management quality metric is established that is designed as a proxy for the 'M' in the CAMEL rating.

For failure prediction, the management quality metric is combined with variables representing the other four factors in the CAMEL rating, as well as a proxy for local economic conditions. The resultant probit-regression model is compared with the leading alternative approaches and evaluated for sensitivity to prediction error. The empirical testing not only shows the new model to be dramatically superior to all other approaches, but underscores the validity and importance of the management quality metric in forecasting. This study not only advances the state-of-the-art in institutional failure prediction, but pioneers the use of DEA as a predictive factor that has far-reaching applications.

# 1 The Role of Management in Banking

Clearly, the quality of a bank's management is key to its long-run survival, and one of management's greatest challenges is coping with the industry's increasing uncertainties and accompanying risks. Internally, this means that bankers must effectively allocate scarce resources, implement controls and procedures to minimize risks and control costs, and be open to the use of new technologies to increase operating efficiencies. Externally, they must keep pace with new regulatory actions, economic fluctuations, societal trends, technological advances, and changes taking place in the global economy.

Kaufman [13] postulated that "to survive in a risky world, banking firms must cope with risk and manage it." This management of risks seems to be the key factor in bank failure. Consider the following statements from recent research studies:

> "...apart from fraud, a single event rarely cripples a bank fatally. Instead, the culprit is consistent mismanagement or risky strategies."[6]
>
> "...bank failure, which affects only a handful of banks, is caused by mismanagement; mismanagement, furthermore, of the basic, old-fashioned risks of banking like lending, liquidity, and controls." [9]
>
> "Management factors generally explain why one bank survives while another fails when facing almost identical circumstances." [7]
>
> "It is the management of the bank that determines success or failure." [18]
>
> "...management incompetence in its broad sense is a major cause of bank failure." [15]
>
> "...the ultimate determinant of whether or not a bank fails is the ability of its management to operate the institution efficiently and to evaluate and manage risk." [20]

Many studies of bank performance and bank failure cite management quality as the most important factor to long-run survival. For a bank to continue to survive, its management must understand, manage, and control the increased risks inherent in today's financial environment. This means that bankers must effectively allocate resources and efficiently control the bank's operations.

# 2 How Can Management Quality Be Measured?

In 1993, Barr, Seiford, and Siems [2] (BSS) presented an approach to measuring a bank's managerial quality using only its financial statements. Their model considers the essential financial intermediation functions of a bank and computes a scalar measure of efficiency. This measure of transformational efficiency is considered a proxy for management quality and is derived using a multiple-input, multiple-output model as follows.

## 2.1 Data Envelopment Analysis: an Overview

In standard microeconomic theory, an enterprise can be viewed as a *decision-making unit* (DMU) that is concerned with transforming a set of $m$ different inputs into $s$ different outputs. For example, BSS models a bank as a DMU with six inputs (full-time-equivalent employees, salary expense, premises and fixed assets, other noninterest expense, total interest expense, and purchased funds) and three outputs (core deposits, earning assets, and total interest income). Decision units of the same kind can be compared for their relative transformational efficiency—this is often performed in traditional analyses using a series of simple ratios of individual outputs to individual inputs.

DEA is a non-parametric method initiated by Charnes, Cooper and Rhodes[10], which uses a generalized ratio to consider multiple inputs and outputs simultaneously. Specifically, if DMU $k$ consumes amounts $X_{ik}$ of inputs, $i = 1, \ldots, m$, and produces amounts $Y_{jk}$ of outputs, $j = 1, \ldots, s$, then its overall efficiency can be given as: $E_k = \sum_{j=1}^{s} Y_{rk} u_{rk} / \sum_{i=1}^{m} X_{ik} v_{ik}$ where $u_{jk}$ and $v_{ik}$ are positive scaling factors, or *weights*, that are to be determined by the DEA process. Rather than using preset weights to evaluate all banks, an optimal set of weights is computed for each bank that makes its ratio as large as possible. These weights are found for an individual bank $k$ by solving the following fractional linear programming problem:

$$
\begin{aligned}
\text{Maximize } & E_k \\
\text{subject to: } E_d & \leq 1, d = 1, \ldots, n \\
u_j, v_i & \geq 0, i = 1, \ldots, m; j = 1, \ldots, s
\end{aligned}
$$

By solving the above problem for each decision unit, an empirical production surface (or efficient frontier) is constructed that shows the maximum amount of outputs that can be achieved by combining various quantities of the inputs, based on the observed behavior of a population. DEA uses linear programming with observed data to locate these frontiers and determine the efficiency of each organizational unit. It determines which of the $n$ DMUs define the surface and, hence are most efficient. Efficient DMUs lie on the surface and have a relative efficiency measure of 1; inefficient DMUs are not on the surface, have an $E_k < 1$.

Note that, unlike regression analysis, DEA is not limited to a single output, and is directed to identify the "best practice" units, not central location of the data. See [23] for a survey of DEA methods, and [22] for a 500-entry bibliography on DEA. Standard computer programs are available to solve such models (see [11]).

## 2.2 Using DEA as a Management Quality Metric

As described above, BSS model a bank as a transformer of six inputs into three outputs. Since such a process reflects the key activities of management, the DEA efficiency metric was expected to be a good proxy for managerial quality and empirical evidence in the BSS study supported this view. Specifically, the research revealed that significant statistical differences in average management quality scores existed up to three years prior to failure. Figure 4 illustrates the study's finding that from December 1984 to June 1987 the average DEA scores for a sample of surviving banks ranged between 0.8142 and 0.8315; whereas, for banks which failed the first half of 1988, the average scores dropped from 0.7986 in December 1984 to 0.6694 in June 1987.

Furthermore, when pooling the data to compute average scores for one year prior to failure, 94.6% of the survivors had scores above 0.7, but only 50% of the failed banks. Additionally, over 50 percent of the survivors scored above 0.8, while only 13.9 percent of the failures did. The average DEA score for failures one year prior to failure was 0.7032 and for survivors it was 0.8203.

Similarly, for banks that are two years prior to failure, there is again a demarcation between surviving banks and failed institutions. However, the differentiation in DEA scores is not as pronounced. For failures, the average DEA score two years prior to failure was 0.7637 and for survivors it was 0.8394.

The empirical results of the BSS study "confirm that the quality of management is crucial to a bank's survival. Scores for surviving institutions are statistically higher than the scores for failed banks. Furthermore, banks that are nearer failure are found to have lower efficiency scores." These results are significant in that banks that survive can be statistically differentiated from banks that fail based on the management quality scores generated by the DEA model. Hence, the use of these scores as a variable in a bank-failure prediction model seems promising.

## 3 Early-Warning Models

An early-warning model is simply an established procedure (usually statistical) for classifying banks into groups (usually *failure* and *non-failure*). This is typically done using only financial characteristics of candidate institutions. The goal of an early-warning model is to identify an institution's financial weakness at the initial stage in its process of deterioration so as to warn interested parties of its potential failure. Banking regulators have recognized the usefulness of early-warning models and have conducted intensive research efforts in this area for nearly two decades so that they can classify problem institutions before they have a major financial impact on the federal deposit insurance fund.

Modern studies of institutional failure began with Beaver [4] who used financial ratios to predict bankruptcy of non-financial firms; this was also the first model to forecast solvency from accounting data . Altman [1] and Sinkey [25] subsequently developed discriminant models for predicting corporate bankruptcy.

Early-warning models for banks began with Meyer and Pifer [17]; since then, various researchers have used multivariate techniques to explain past closures and predict future failures. While discriminant models have evolved over the years, the variables used to predict failure have largely remained constant. For financial institutions, they generally fall into the CA_EL categories and, more recently, included variables to capture operating efficiencies and local economic conditions. (For other readings on early-warning models, see [8, 12, 17, 21, 26].)

Generally, the measures used to evaluate operating efficiencies are regarded as proxy variables for the quality of management. The major complication, however, is that it takes many ratios to encompass the operating characteristics of an organization. Hence, to measure management quality, most recent studies have focused on the ratio of total operating income to total operating expense.

Three prominent models are compared with the new models developed in this study: those appearing in [12, 16, 18]. These were selected because they use economic and publicly-available financial data, none of which require on-site examination. While all three studies did not define "failure" the same way, the traditional view that failure occurs through a declaration of insolvency by one of the regulatory agencies is adhered to for this test. All models utilize probit or logit regression as the statistical classification technique and the data used for the survivors in each analysis were randomly selected. Table 1 provides a comparison of the three models using information presented in their respective articles.

These three models have many other similarities: (1) all reflect capital adequacy with a single ratio of capital to assets; (2) each contains two asset quality variables; and (3) each represents earnings with a net-income-to-total-assets ratio. (Hanweck includes an additional earnings variable to capture the percentage change in net operating income.) The models are differentiated by the fact that Hanweck included an institutional scale variable and Pantalone and Platt included a local economic conditions term.

Martin [16], 1970-76 Data: 5642 Survivors, 58 Failures

| *Logit Model Variables* | *Coeff* | *Sig*[a] |
|---|---|---|
| Constant | -5.33 | ** |
| Net income / total assets | -120.86 | ** |
| Gross charge-offs / net operating income | 2.20 | ** |
| Commercial & industrial loans / total loans | 7.89 | ** |
| Gross capital / risk assets | -35.63 | ** |

Hanweck [12], 1973-75 Data: 177 Survivors, 32 Failures

| *Probit Model Variables* | *Coeff* | |
|---|---|---|
| Constant | -4.14 | * |
| Net operating income / total assets | -69.49 | ** |
| Equity capital / total assets | 14.86 | |
| % change in net operating income / total assets | -0.01 | |
| % change in total assets | -1.18 | |
| Loans / capital | 0.26 | ** |
| Size: log(assets) | 0.02 | |

Pantalone & Platt [18], 1983-84 Data: 226 Survivors, 113 Failures

| *Logit Model Variables* | *Coeff* | |
|---|---|---|
| Constant | -0.01 | |
| Net income / total assets | -71.39 | ** |
| Equity capital / total assets | -11.79 | ** |
| Total loans / total assets | 7.71 | ** |
| Commercial & industrial loans / total loans | 3.72 | ** |
| \|%\| change in residential construction | 0.10 | |

[a] Coefficient is significant: * = at the 0.05 level; ** = at the 0.01 level

Table 1: Summary of the Test Set of Bank-Failure-Prediction Models

# 4 Developing New Early-Warning Models to Predict Bank Failure

For an early-warning model to be reliable and useful, other facets of a bank should be analyzed in harmony with the quality of management. To date, the most effective early-warning models attempt to replicate the regulators' CAMEL rating scheme. Thus, a model that identifies variables to represent the factors in the CAMEL rating should not only prove to be among the most accurate, but have a greater chance of acceptance by regulatory bodies. Additionally, since it appears that local economic conditions play an important role in the success or failure of financial institutions, a variable to proxy a bank's local economic conditions is also considered.

The models in this paper are developed from the regulator's viewpoint and designed to operate on current data. In other words, for a model devised to predict failure $t$ periods ahead, the variables considered in the model must be available at least $t$ periods prior to the time of failure.

The new models embody two forecast lead-times: prediction 12 to 18 months prior to failure (referred to as a "one-year-ahead" model) and 24 to 30 months prior (referred to as a "two-year-ahead" model). For both lead-times, models based on the probit regression classification methodology are implemented.

We hypothesize that the quality of management is important to a bank's survival. Hence, the DEA efficiency measure should be a highly significant variable in the failure-prediction equation. We expect that our new model will classify survivors and failures more accurately when the management quality metric is included, and will out-perform previously published classification schemes.

## 4.1 Data Sets Used for Model Construction and Validation

For both the one-year-ahead (1YA) and two-year-ahead (2YA) models, commercial banks which failed from 1986 through 1988 and a random sample of banks which survived through 1989 composed the population. The following restrictions were in place: (1) banks had to have been in operation for at least three years, and (2) only banks with total asset size of between \$20 and \$300 million were included.

The age limitation was intended to keep *de novo* institutions from entering into the failure equations. New institutions often have anomalous balance sheets and income statements that could negatively affect the reliability and usefulness of a failure-prediction model. Also, since the regulatory agencies charter new financial institutions, they already have controls in place to meet their goal of fostering a safe and sound banking system by restricting entry to the industry to only those with strong prospects for success.

The size limitations were invoked because of the tremendous differences in the managerial operations between small, medium-sized, and large banks. The largest banking institutions (i.e., those with over $300 million in total assets) were eliminated from the sample population because they are generally watched more closely and examined more frequently by banking regulators. In general, the larger the bank, the more frequent and thorough the examination. In fact, it is common for bank examiners to be occupied in an on-site examination of a large bank for months, sometimes years. With near-continuous monitoring, there is no real need for an early-warning model.

The smallest banks (those under $20 million in assets) were dropped from the population for the following reasons.

- Managing a small bank is much different than managing a larger bank. Research has identified major differences in performance, markets, and operating costs (see [5, 14, 19]). Smaller banks have internal operating disadvantages, primarily related to the absence of economies of scale.
- These smaller banks often behave different from "traditional" banks. By accepting deposits and buying securities, some act as investment houses or mutual funds. Others use borrowed money predominantly to make loans, relying little on core deposits.
- The deposits for the smallest banks are small relative to larger institutions: only 1.3% of total banking deposits. One could argue that on-site examinations may not be worth their cost and effort for the smallest institutions.
- As banks grow through increased mergers and acquisitions, the smallest banks are gradually disappearing, while the percentage of medium-sized and large banks has increased.

The population used in this study includes over 70% of the financial institutions in operation in the United States, and most of the "at risk" banks. Economic and financial statement data were employed for each six-month period beginning with December 31, 1984 and ending with June 30, 1988.

As shown in Table 2, the data for failing institutions were divided into eight time periods, based on their failure date. Thus, if failure occurred during the last six months of 1988, economic and financial figures for June 30, 1987 were used as the one-year-ahead data. Data for time periods 1 through 6 were then pooled and, together with data from the survivors, used to build each model.

Data for the survivors were divided into the same eight time periods and grouped so that no bank would be represented in two different time periods. In other words, of the 611 survivors in the population, banks were sampled at random without replacement and placed into one of the eight time periods. This grouping was done to ensure that the data for the surviving banks would be independent and identically distributed. This random selection scheme reduced

| Time Period | Statement Date | Failure Date | No. of Survivors | No. of Failures | Total Banks |
|---|---|---|---|---|---|
| 1 | Dec. 31, 1984 | 1/86–6/86 | 71 | 40 | 111 |
| 2 | June 30, 1985 | 7/86–12/86 | 74 | 46 | 120 |
| 3 | Dec. 31, 1985 | 1/87–6/87 | 75 | 40 | 115 |
| 4 | June 30, 1986 | 7/87–12/87 | 74 | 48 | 122 |
| 5 | Dec. 31, 1986 | 1/88–6/88 | 75 | 51 | 126 |
| 6 | June 30, 1987 | 7/88–12/88 | 76 | 69 | 145 |
| | **Model Construction Totals:** | | **445** | **294** | **739** |
| | | | | | |
| 7 | Dec. 31, 1987 | 1/89-6/89 | 76 | 54 | 130 |
| 8 | June 30, 1988 | 7/89–12/89 | 76 | 65 | 141 |
| | **Holdout Sample Totals:** | | **152** | **119** | **271** |

Table 2: Summary of Test Data for One-Year-Ahead Models

the total number of survivors represented in the model to 597 because of missing observations for some of the banks. (Due to the size boundaries, some banks move into and out of the population at various dates in their time series. This could result in a bank not being represented in the data used to build and test the model. Because this predicament exists when using the model in practice, the slight reduction in the number of survivors represented in the population was not seen as a problem.)

To develop the 1YA failure-prediction model, the data were aggregated for time periods 1 through 6. Hence, there were 445 survivors and 294 failures in the population with all observations independent and identically distributed random variables. The holdout sample consisted of time periods 7 and 8, with a total of 152 survivors and 119 failures.For the 2YA model, a similar selection process was used, but it resulted in a slightly different number of banks in the population. To construct the 2YA model, there were 391 survivors and 211 failures. The holdout sample consisted of 202 survivors and 126 failures.

## 4.2 Variable Selection

Our new models include only six variables: one for each factor of the CAMEL rating and another to proxy local economic conditions. This restriction was designed to keep them parsimonious and understandable to regulators and bankers. A large number of variables were analyzed statistically (see Barr and Siems [3] for details] before selecting those described in Table 3.

While some of the variables have appeared in earlier models, this is the first study in which the DEA efficiency score has been used in a bank forecasting model, or any regression model in the literature. All values—including DEA—

| Variable | One Year Ahead | | Two Years Ahead | |
|---|---|---|---|---|
| | Survivors | Failures | Survivors | Failures |
| Equity Capital/Total Loans | 18.85% | 8.56% | 18.65% | 10.83% |
| Nonperforming Loans/Total Assets | 1.42% | 5.63% | 1.28% | 3.27% |
| DEA Efficiency Score | 0.8203 | 0.7033 | 0.8394 | 0.7637 |
| Net Income/Total Assets | 0.89% | -2.18% | 0.93% | -0.62% |
| Large Deposits/Total Assets | 10.03% | 24.97% | 9.98% | 26.12% |
| % Change In Residential Construction | -1.68% | -15.55% | 0.33% | -12.55% |

Table 3: Comparison of Variable Averages for One-Year and Two-Years Prior To Failure

are based on publicly available data, and many are from the Federal Reserve System's Call Report Database. Table 3 shows, for each variable included in the 1YA and 2YA models, the mean and standard deviation for both the survivors and failures. Also given are the results of a $t$-test on each variable, testing the hypotheses that the true survivor and failure means are equal: the $t$-statistic and the significance level, or probability of identical means.

All of the selected variables show significant differences in means between the survivor and failure groups, in both 1YA and 2YA cases. All of the selected variables have significantly different means for survivors versus failures and the differences in mean values follow natural expectations: survivors had higher capital, better quality assets, more efficient management, higher earnings, and more liquidity.

In summary, the six variables selected for our failure-prediction models were:

| *Variable* | *Proxy for* |
|---|---|
| Equity Capital/Total Loans | Capital Adequacy |
| Non-performing Loans/Total Assets | Asset Quality |
| DEA Efficiency Score | Management Quality |
| Net Income/Total Assets | Earnings Ability |
| Large Dollar Deposits/Total Assets | Liquidity |
| Percentage Change in Residential Construction | Local Economic Conditions |

A comparison of the 1YA and 2YA mean values gives insight into differences between the surviving and failing populations. Except for the liquidity measure, the gap between the mean values for survivors and failures widens as failure approaches. Also, there is little difference in mean values for the survivors, but a noticeable deterioration for failures (except for the liquidity variable). Thus, one might expect the 1YA model to predict failure with greater accuracy than the 2YA model.

| | One Year Ahead (LR Index: .72) | | Two Years Ahead (LR Index: .58) | |
|---|---|---|---|---|
| Variable | Coeff. | t-ratio | Coeff. | t-ratio |
| Constant | 5.1395 | 6.747† | 2.9100 | 3.901† |
| Equity Capital/Total Loans | -9.6993 | -4.982† | -9.4078 | -4.854† |
| DEA Efficiency Score | -7.7682 | -8.457† | -4.6943 | -5.500† |
| Nonperforming Loans/ Total Assets | 17.8065 | 5.442† | 24.5388 | 4.990† |
| Net Income/Total Assets | -22.0646 | -4.080† | -3.9871 | -.635† |
| Large Deposits/Total Assets | 5.8907 | 7.060† | 5.9031 | 8.209† |
| % Change in Residential Construction | -2.7024 | -4.567† | -3.0982 | -4.850† |
| Percent Correctly Classified | In-Sample | Out-Sample | In-Sample | Out-Sample |
| Survivors | 94.4% | 96.1% | 92.1% | 92.1% |
| Failure | 89.5% | 96.6% | 82.9% | 94.4% |
| Total | 92.4% | 96.3% | 88.9% | 93.0% |

†Significant at the 0.01 level.

Table 4: Full Models and Classification Results

### 4.3 Empirical Testing: The One-Year-Ahead Model

The standard probit methodology was used to develop models that would classify banks as either survivors or failures. The models were configured so that the dependent variable took on the value of zero (0) for survivors and one (1) for failures. Hence, a negative (positive) coefficient means that the variable is inversely (directly) related to failure and directly (inversely) related to survival. When used for classification, a fractional *cut-off value* for the computed dependent variable—usually 0.5—assigns an observation to either the survivor or failure group.

The 1YA model, with the coefficients and $t$-ratios, is presented in Table 4. To develop this model (and others in the study), a pooled time-series cross-sectional dataset was constructed. Tests for aggregation bias concluded that a structural shift in the data was detected in time period 6. However, because the shift did not classify banks better than the holdout sample, the structural shift model was discarded.

For this model, all variables were significant at the 0.01 level and the signs on the variables' coefficients are appropriate. Specifically, for the variables with positive coefficients—ratios of non-performing loans to total assets, and large dollar deposits to total assets—one would expect that the higher the variable level, the higher the probability of failure (because the expected score becomes larger). Similarly, for the variables with negative coefficients—equity-capital-to-total-loans ratio, the DEA efficiency score,ratio of net income to total assets, and the percentage change in residential construction—one would expect that the higher the variable level, the lower the probability of failure.

The variable with the highest $t$-ratio, in absolute value terms, was the DEA efficiency score with a $t$-ratio of -8.457. Thus, it appears that the quality of management added the most information to the new 1YA bank-failure prediction model.

The *likelihood ratio index* (identified as the LR index in the table) measures the goodness of fit as a pseudo-$r^2$. For this model the likelihood ratio index was 0.72—a much better fit than the Martin [16], Hanweck [12], and Pantalone and Platt [18] models with LR indices of 0.42, 0.40, and 0.51, respectively, on the same dataset.

The classification results for this new 1YA bank failure-prediction model were exceptionally accurate. As shown in the lower portion of Table 4, using a classification cut-off of 0.5, the model correctly identified 94.4% of the survivors and 89.5% of the failures for a total classification accuracy of 92.4% (the total number of correct classifications divided by the total number of banks in the sample).

The results for the holdout sample were even stronger. For the two holdout time periods combined, the new model correctly classified 96.1% of the survivors and 96.6% of the failures for an overall classification accuracy of 96.3%. The classification results for all of the models discussed in this paper were statistically significant at the 0.01 level as compared to the results of a naive forecast based on sample proportions.

### 4.4 Empirical Testing: The Two-Year-Ahead Model

For the models which predict bank failure 24 to 30 months ahead, the proxy variable for earnings was no longer significant, as shown in Table 4. Additionally, the most significant variable (in terms of greatest absolute t-ratio) is no longer the management quality term but, rather, liquidity: large-dollar deposits over total assets.

The 2YA model had an LR index of roughly 0.58 compared to 0.72 for the 1YA model. This is lower because by using data that are further from the actual failure date, the two groups (survivors and failures) tend to look more similar.

Also, as expected, the classification results of the 2YA model were not as accurate as those for the 1YA model. Using a classification cut-off of 0.5, the percentage of correctly identified banks for the in-sample group was 88.9%, with 92.1% of the survivors and 82.9% of the failures correctly classified. However for the holdout sample, the 2YA model correctly identified 92.1% of the survivors, and 94.4% of the failures, for a strong 93.0% overall accuracy.

| | One Year Ahead (LR Index: .62) | | Two Years Ahead (LR Index: .54) | |
|---|---|---|---|---|
| Variable | Coeff. | t-ratio | Coeff. | t-ratio |
| Constant | -.8165 | -3.037† | -.7913 | -2.515† |
| Equity Capital/Total Loans | -9.4796 | -5.575† | -9.8316 | -5.369† |
| Nonperforming Loans/ Total Assets | 19.9399 | 6.747† | 23.2073 | 5.114† |
| Net Income/Total Assets | -18.1565 | -3.931† | -2.8936 | -.490† |
| Large Deposits/Total Assets | 5.4826 | 7.752† | 6.0231 | 8.785† |
| % Change in Residential Construction | -2.3222 | -4.520† | -3.1322 | -5.247† |
| Percent Correctly Classified | In-Sample | Out-Sample | In-Sample | Out-Sample |
| Survivors | 92.4% | 96.1% | 91.3% | 90.6% |
| Failure | 84.0% | 93.3% | 80.6% | 95.2% |
| Total | 89.0% | 94.8% | 87.5% | 92.4% |

†Significant at the 0.01 level.

Table 5: Model Results without the Management Quality Metric

### 4.5 Effect of Removing the Management Quality Variable

Of concern in this study is the impact of the DEA efficiency score on the model. Since it is the most significant variable in the 1YA instance, one might expect that the fit and classification results would deteriorate if this variable were eliminated.

To test the impact of the management quality variable, it was removed from the models and a second analysis performed. The coefficients and classification result are given in Table 5 for the "reduced" 1YA and 2YA models. For the 1YA case, the original variables all remain significant with the expected signs. The LR index of fit drops from 0.72 for the full model to 0.62 for the reduced model, and the classification accuracy is lower. With a 0.5 classification cut-off, 89.0% of all in-sample banks were correctly classified and the out-of-sample accuracy dropped from 96.3% to 94.8%. Removing the DEA efficiency variable from the 2YA model has a similar effect: the significance of each variable remains roughly the same as before, the fit metric drops, and the classification results are lower.

These results demonstrate the importance of bank management in the success or failure of a bank. The models which included the management quality variable fit the original data much better and classified failures and survivors more accurately for both in-sample and holdout-sample tests.

| Model | LR index | In-Sample Correct Classification Survi-vors % | Fail-ures % | Total % | Out-of-Sample Correct Classification Survi-vors % | Fail-ures % | Total % |
|---|---|---|---|---|---|---|---|
| One Year Ahead: | | | | | | | |
| Martin | .42 | 93.0 | 71.1 | 84.3 | 89.5 | 94.1 | 91.5 |
| Hanweck | .40 | 92.8 | 70.1 | 83.8 | 94.7 | 86.6 | 91.1 |
| Panalone & Platt | .51 | 89.7 | 77.9 | 85.0 | 90.1 | 90.8 | 90.4 |
| New model w/o DEA | .62 | 92.4 | 84.0 | 89.0 | **96.1** | 93.3 | 94.8 |
| New model with DEA | **.72** | **94.4** | **89.5** | **92.4** | **96.1** | **96.6** | **96.3** |
| Two Years Ahead: | | | | | | | |
| Martin | .22 | **94.1** | 46.0 | 77.2 | 96.5 | 61.9 | 83.2 |
| Hanweck | .18 | 93.9 | 41.7 | 75.6 | **97.0** | 54.8 | 80.8 |
| Panalne & Platt | .38 | 89.8 | 71.1 | 83.2 | 92.6 | 84.1 | 89.3 |
| New model w/o DEA | .54 | 91.3 | 80.6 | 87.5 | 90.6 | **95.2** | 92.4 |
| New model with DEA | **.58** | 92.1 | **82.9** | **88.9** | 92.1 | 94.4 | **93.0** |

Table 6: Comparative Classification Accuracy Results for All Models

### 4.6 Classification Accuracy Analysis

Table 6 summarizes the classification results on our data sets for the Martin, Hanweck , Pantalone and Platt, and all of our new models (with and without the DEA variable). For each model, the table shows the LR index of fit, in-sample classification accuracy rates, and holdout sample classification accuracy rates for both survivors and failures. All classifications are based on a 0.5 logit/probit cut-off point (discussed further in the next section).

For the 1YA data, the two new models classified both failures and survivors more accurately than the previously developed models from the literature. In addition, the model which includes the management quality variable was the most accurate of all.

Similarly, classification accuracy comparisons are shown for the 2YA models. Again, the new bank-failure prediction model which incorporates the management quality variable has the best fit and classified banks more accurately both in-sample and out-of-sample.

To demonstrate the robustness of the new models, Figure 1 compare the overall in-sample classification accuracy rates at cut-off points ranging from 0.1 to 0.9 for all of the models for the 1YA and 2YA models, respectively. In both cases, our models that include the managerial quality variable out-perform all of the other models analyzed for overall accuracy. In fact, the worst case for our 1YA model is more accurate than the best cases of the Hanweck, Martin,

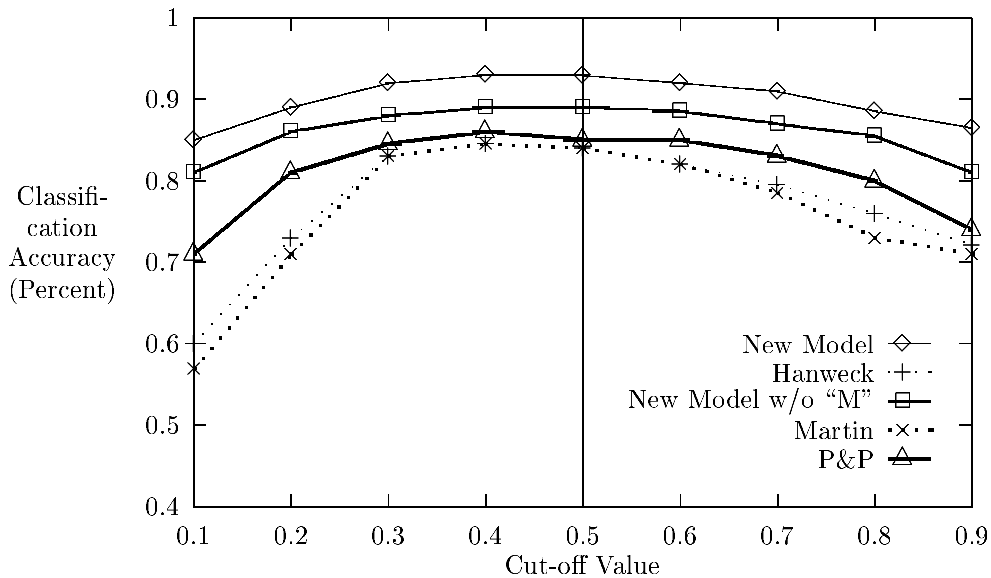


Figure 1: Effect of the cutoff value on the 1YA models' classification accuracy.

and Pantalone and Platt models.

## 5 Costs of Misclassifications

Most published studies assert that the cost of misclassifying a bank that fails (a Type I error) is greater than the cost of misclassifying a bank that continues to survive (a Type II error). The reasoning is that the cost to perform an on-site examination that results in significant operating improvements is less than the cost to bail out the same bank if it was not examined and had failed. Although this is the conventional wisdom, researchers have not identified the actual costs of misclassifications to permit minimization of the total expected costs.

The likelihood of both types of errors is affected by the selection of the probit cut-off point. A lower cut-off will decrease the probability of a Type I error, while increasing the likelihood of a Type II. Table 7 shows the number and type of in-sample misclassifications for different cut-off points applied to the 1YA model. Hence most researchers have taken the stance, advocated by Pantalone and Platt, that a cut-off less than 0.5 should be used but have not justified their chosen value.

The "best" probit classification cut-off point minimizes the total cost of misclassifying banks. We analyzed four different scenarios, where the ratios

| Cut-off Point | Number of Misclassifications | | Relative Cost Magnitude Type I Error:Type II Error | | | |
|---|---|---|---|---|---|---|
| | Type I | Type II | 1:1 | 2:1 | 5:1 | 10:1 |
| 0.1 | 7 | 104 | 111 | 118 | 139 | **174** |
| 0.2 | 12 | 66 | 78 | 90 | 126 | 186 |
| 0.3 | 15 | 45 | 60 | **75** | **120** | 195 |
| 0.4 | 23 | 31 | **54** | 77 | 146 | 261 |
| 0.5 | 31 | 25 | 56 | 87 | 180 | 335 |

Table 7: Relative Costs Of Misclassifications for the One-Year-Ahead Model

of Type I error costs to Type II error costs are: 1:1, 2:1, 5:1, and 10:1. For example, the 10:1 ratio assumes the cost of misclassifying a bank that in fact fails is ten times the cost of misclassifying a bank that survives; 1:1 assumes that the two costs are equal.

Shown in Table 7 is the total misclassification cost, defined as the relative cost of a Type I error multiplied by the number of failures misclassified, plus the cost of a Type II error multiplied by the number of survivors misclassified. For the 1:1 scenario, the total cost is the sum of the number of banks misclassified. When misclassifying failures is twice as costly as examining survivors (column 2:1), the total cost is equal to twice the number of failures misclassified (Type I error) plus the number of survivors misclassified (Type II error).

Under the four cost magnitudes, a minimum value can be found that corresponds to the cut-off value that should be chosen. For the 1:1 case, a cut-off of 0.4 will yield the minimum total cost of 54. For the 2:1 and 5:1 scenarios, a cut-off point of 0.3 should be used, and a value of 0.1 should be used for the 10:1 scenario. Search procedures could be applied to find the optimal cut-off point.

# 6 Policy Implications

More accurate bank failure-prediction models, such as those developed in this study, have four key implications for policy makers. First, more accurate models allow banking regulators to deploy their examination resources more efficiently. An effective early-warning model will detect and classify the weakest financial institutions. With banks rated on a survival/failure likelihood scale, regulators can focus on those that are the greatest threat to the deposit insurance fund. The impacts not only include a reduction in expenses for the regulatory agencies (through more efficient resource allocation) but a potential savings to the taxpayers through early identification of weak banks that can be "turned around" before threatening the deposit insurance fund.

Second, the models developed in this study can be used by regulators and bankers to gain a better understanding of the reasons for bank failure. The 1YA model shows that management quality has the greatest impact on bank failure; banks receiving the highest DEA efficiency scores are much more likely to survive than banks which have relatively low scores. The 1YA model identifies five additional variables that are significant in differentiating between failures and survivors: low capital (as a percentage of loans), a high level of non-performing loans, low net income, a high level of large dollar deposits, and slow growth (or rapid decline) in the annual rate of residential construction in the bank's home state.

If regulators have a better understanding of why banks fail, they will be better able to offer bankers suggestions on how to become more stable. In addition, managers will be alerted to threatening conditions so that action can be taken before more serious problems arise.

A better bank failure-prediction model strengthens the entire examination process by identifying banks more objectively. Bankers can have a better understanding when examiners suggest changes in their operations. Additionally, the value of the quarterly call report data would be greatly enhanced.

Finally, the new early-warning models developed in this study could be used to develop a variable-rate deposit-insurance-premium structure. The banks at the greatest risk of failure would be required to pay the highest premiums for deposit insurance, much like smokers pay higher life insurance premiums than nonsmokers.

# 7 Conclusions

In this paper, two new bank-failure prediction models were developed. Both the one-year-ahead and two-year-ahead models use proxy variables for each factor in the CAMEL rating plus a variable to capture local economic conditions. For the first time, bank failure-prediction models were developed using a DEA efficiency variable to proxy management quality.

The results of these models indicate that management is, indeed, important to the successful operation of a bank. When the management variable was removed from the full model, the results were worse in terms of the model's fit to the data and its classification accuracy. The newly-developed models also show superior results to leading published approaches, as analyzed using a standardized dataset.

From a policy standpoint, these early-warning models can be used to detect and classify the weakest financial institutions, while providing regulators and bankers insight into the reasons for bank failure. Their high accuracy should allow banking regulators to allocate more efficiently their resources to perform on-site examinations, and should strengthen the entire examination process by providing a simple and objective detection mechanism for failure-prone banks.

Such new tools will ably assist those striving to ensure a safe and secure future for the nation's banking system.

# References


[1] Altman, E. I. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *The Journal of Finance,* 1968, *23,* 589-609.

[2] Barr, R. S., Seiford, L. M., and Siems, T. F. An Envelopment-Analysis Approach to Measuring the Management Quality of Banks, *Annals of Operations Research*, 1993, *38*, 1-13.

[3] Barr, R.S. and Siems, T.F. Predicting Bank Failure Using DEA to Quantify Management Quality, working paper, Federal Reserve Bank of Dallas (Dallas, TX), 1993.

[4] Beaver, W. H. Financial Ratios as Predictors of Failure, Empirical Research in *Accounting: Selected Studies* (The Institute of Professional Accounting, University of Chicago), 1966, 71-127.

[5] Bentson, G. J., Hanweck, G. A., and Humphrey, D. B. Operating Costs in Commercial Banking, *Economic Review*, 1982, November, 6-21.

[6] Booker, I. O.Tracking Banks from Afar: A Risk Monitoring System, *Economic Review*, 1983, November, 36-41.

[7] Bovenzi, J. F., and Nejezchleb, L. Bank Failures: Why Are There So Many?, *Issues in Bank Regulation*, 1985, *8,* 54-68.

[8] Bovenzi, J. F., Marino, J. A., and McFadden, F. E. Commercial Bank Failure Prediction Models, *Economic Review*, 1983, *68*, November, 14-26.

[9] Cates, D. C. Bank Risk and Predicting Bank Failure,*Issues in Bank Regulation*, 1985, *9*, 16-20.

[10] Charnes, A., Cooper, W.W., and Rhodes, E. Measuring the Efficiency of Decision Making Units,*European Journal of Operational Research*, 1978, *2*, 429-444.

[11] Charnes, A., Cooper, W.W., Lewin, A.Y., and Seiford, L. *Data Envelopment Analysis: Theory, Methodology and Applications*, 1994, Kluwer Academic Publishers, Boston.

[12] Hanweck, G. A. Predicting Bank Failure, *Research Papers in Banking and Financial Economics*, 1977, (Financial Studies Section, Division of Research and Statistics, Board of Governors of the Federal Reserve System), November.

[13] Kaufman, G. G. Banking Risk in Historical Perspective, *Bank Structure and Competition*, 1986, May, 231-249.

[14] King, B. F. Changes in Large Banks' Market Shares, *Economic Review*, 1982, November, 35-40.

[15] Looney, S. W.,Wansley, J. W., and Lane, W. R. An Examination of Misclassifications with Bank Failure Prediction Models, *Journal of Economics and Business*, 1989, *41*, 327-336.

[16] Martin, D. Early-warning of Bank Failure: A Logit Regression Approach, *Journal of Banking and Finance*, 1977, *1*, 249-276.

[17] Meyer, P. A., and Pifer, H. W. Prediction of Bank Failures, *The Journal of Finance*, 1970, *25*, 853-868.

[18] Pantalone, C. C., and Platt, M. B. Predicting Commercial Bank Failure Since Deregulation, *New England Economic Review*, 1987, Jul/Aug, 37-47.

[19] Rhoades, S. A., and D. T. Savage Post-Deregulation Performance of Large and Small Banks,*Issues in Bank Regulation,* 1990, Winter, 20-31.

[20] Seballos, L. D. and Thomson, J. B. Underlying Causes of Commercial Bank Failures in the 1980s, *Economic Commentary*, 1990, Federal Reserve Bank of Cleveland, September.

[21] Secrist, H.*National Bank Failures and Non-Failures: An Autopsy and Diagnosis*, 1938, Principia Press, Bloomington, IN.

[22] Seiford, L.M. A Data Envelopment Analysis Bibliography (1978-1992), in nes, A., Cooper, W.W., Lewin, A.Y., and Seiford, L. *Data Envelopment Analysis: Theory, Methodology and Applications*, 1994, Kluwer Academic Publishers, Boston.

[23] Seiford, L.M. and Thrall, R.M. Recent Developments in DEA: The Mathematical Approach to Frontier Analysis,*Journal of Econometrics*, 1990,*46*, 7-38.

[24] Siems, T. F. Quantifying Management's Role in Bank Survival, *Economic Review*, 1992, First Quarter, 29-41.

[25] Sinkey, J. F., Jr. *Problem and Failed Institutions in the Commercial Banking Industry*, 1979, JAI Press, Greenwich, CT.

[26] West, R. C. A Factor-Analytic Approach to Bank Condition, *Journal of Banking and Finance*, 1985,*9*, 253-266.