Predicting Large U.S. Commercial Bank Failures

James Kolari, Dennis Glennon, Hwan Shin, and Michele Caputo

Economic and Policy Analysis Working Paper 2000-1

January 2000

Predicting Large U.S. Commercial Bank Failures

James Kolari, Dennis Glennon, Hwan Shin, and Michele Caputo*

Abstract: The present study applies empirical methods to the problem of predicting large U.S. commercial bank failures. Because of sampling limitations, scant research has examined the feasibility of using computer-based early warning systems (EWSs) to make such predictions. In the late 1980s and early 1990s numerous large banks failed in the United States, enabling us to collect a sample of 50 failed banks with more than $250 million in assets as well as a matched sample of 50 non-failed large banks. These samples were split into original and holdout samples of different sizes. Both the parametric method of logit analysis and the nonparametric approach of trait recognition are employed to (1) develop classification EWS models based on the original samples and (2) test the predictive ability of these models using the holdout samples. Both logit and trait recognition performed well in terms of classification results. However, over the holdout samples, trait recognition outperformed logit in a variety of tests, including overall accuracy, large bank failure accuracy, weighted efficiency scores, and stability using data from one year before, as well as two years before, failure. Other results from the trait recognition models reveal that complex two- and three-variable interactions between financial and accounting variables contain additional information about bank risk not found in the individual variables themselves. We conclude that nonparametric EWSs can provide valuable information about the future viability of large banks.

---

*The authors are Chase Professor of Finance, Texas A&M University; Senior Economist, Office of the Comptroller of the Currency; Ph.D. Candidate, Texas A&M University; and Professor of Geophysics, Istituto di Fisica della Università di Roma, Rome, Italy, respectively.

The opinions expressed here are those of the authors and do not necessarily represent those of the Office of the Comptroller of the Currency. The authors would like to thank participants in a seminar at the Office of the Comptroller of the Currency for helpful comments. Computer programming assistance by Drew Wagner, Matthew Butler, and especially Rajneesh Khanna was invaluable in creating the trait recognition software for this research. Also, financial support from the Center for International Business Studies at Texas A&M University is gratefully acknowledged. We would like thank Catharine Lemieux, Julapa Jagtiani, Michele Gambera, James Moser, Elijah Brewer, Robert DeYoung, Jeff Brown, Gary Whalen, and others for helpful comments in the course of this research.

Please address correspondence to: Dr. James Kolari, Texas A&M University, Finance Department, College Station, TX 77843 Office Phone: 409/845-4803 Fax: 409-845-3884 Email: j-kolari@tamu.edu
An electronic copy of the trait recognition software and manual is available upon email request from the authors.

I. Introduction

Seminal work by Beaver (1966) and Altman (1968) showed how computer-based models relying on accounting information could predict the failures of firms. Their work sparked a continuing stream of research in the corporate financial literature (e.g., see Beaver (1968), Edmister (1972), Blum (1975), Altman, Haldeman, and Narayanan (1977), Martin (1977), Ohlson (1980), Scott (1981), McFadden (1983), Zavgren (1985), Jones (1987), Keasey and McGuinness (1990), Platt and Platt (1990), Altman (1993), Coats and Fant (1993), Altman, Marco, and Varetto (1994), Altman and Narayanan (1997), and others).[1] Bank regulators are keenly interested in applying these methods to banks to supplement the information they receive from onsite examinations. Computer-based models could be used as early warning systems (EWSs) to help prevent some bank failures or reduce the cost of failure. Extensive research on failed banks has confirmed that computer-based models perform well as EWSs (e.g., see Meyer and Pifer (1970), Stuhr and Van Wicklen (1974), Sinkey (1975), Santomero and Vinso (1977), Bovenzi, Marino, and McFadden (1983), Korobrow and Stuhr (1985), West (1985), Maddala (1986), Lane, Looney, and Wansley (1986), Whalen and Thomson (1988), Espahbodi (1991), Thomson (1993), Kolari, Caputo, and Wagner (1996), and others).

Few studies have sought to determine whether the failures of *large* banks are predictable. Previous work on predicting large bank failures have focused on the usefulness of stock price data as a bank-specific EWS (e.g., see Pettway (1976, 1980), and Peavy and Hempel (1998)). Alternatively, several other studies have attempted to use the financial ratios of individual banks to predict their failure [e.g., see Sinkey (1985) and Federal Deposit Insurance Corporation (1997)]); however, inadequate sample sizes have prevented analysts from studying large banks generically.[2] In this regard, the search for ways to preclude the failures of large banks is becoming increasingly important, because the ongoing consolidation of the banking industry is increasing these banks' numbers (see Berger (1995) and Boyd and Graham (1996)). Such consolidation raises new policy challenges for regulatory and government entities charged with the responsibility of ensuring the banking system's safety and soundness. One policy response to the potential dangers of banks that are too-big-to-fail (TBTF) (including competitive inequalities, moral hazard, and inefficiency) is to increase bank regulation of large institutions (see Hoenig (1999)).[3]

The development of computer-based EWSs for large banks would be consistent with this policy. To create a database for such EWSs, we collected data from the period 1989 to 1992, when numerous large banks failed in the United States. We were able to gather a sample of 50 large failed banks with more than $250 million in assets.[4] Although a total of 50 large bank failures is large by historical standards, it is quite small in terms of minimum sampling requirements in most EWS models. It is common practice to split the sample of failed banks into (1) an original sample that gives rise to a classification model and (2) a holdout sample that is reserved for determining the EWS model's efficacy. In our case this sampling design leaves only 25 large failed banks in the original and holdout samples and only 25 large non-failed banks in the matched-pair (by size and location) original and holdout samples. Both parametric and nonparametric EWS models are tested using these samples. Because of its widespread application in previous finance and banking studies, the parametric approach of logit analysis was chosen. Also, we selected the nonparametric approach of trait recognition as applied to bank failures in Kolari, Caputo, and Wagner (1996) because of its reported usefulness on small samples.[5] *A priori*, we expect that nonparametric EWSs will experience less difficulties with small sample sizes than parametric EWSs because the latter models are likely to violate variable distribution assumptions. To further examine the effects of small samples of large banks on the EWSs, we repeat the comparative analyses using smaller original and holdout samples which have larger minimum asset sizes (e.g., a second sample with the largest 15 failed and 15 non-failed banks and a third sample with the largest 10 failed and 10 non-failed banks in the original and holdout samples).

Using accounting data from both one year and two years prior to the failures, we found that computer-based EWSs are a viable means of evaluating large bank failure risk. Both logit and trait recognition performed well in the classification results of original samples; the accuracy rates were between 90 percent and 100 percent. However, with regard to the prediction results using holdout samples, trait recognition outperformed logit in such tests as overall accuracy, large bank failure accuracy, weighted efficiency scores, and stability using data from one year and two years prior. In regard to the main task of predicting large bank failures, a particularly noteworthy finding is that, while logit predicted large bank failures no better than chance in holdout samples, trait recognition was able to predict most of the large bank failures both one year and two years prior to collapse. Other results from the trait recognition models reveal that complex two- and three-variable interactions between financial and accounting variables contain information about bank risk not found in individual variables. We conclude that nonparametric EWSs can provide valuable information about the future viability of large banks.

The remainder of the paper is organized as follows: Section II overviews logit and trait recognition EWS models, section III describes our empirical methodology, section IV reports the empirical results, and section V provides a summary and conclusion.

II. EWS Models

Parametric Modeling Approach -- Logit Analysis

A logistic distribution is used in many limited dependent variable applications. The resulting model (i.e., a logit model) is common in the EWS literature of finance and banking. The posterior probability of failure can be derived directly from the following logit specification:

$$\log[P_i/(1-P_i)] = a + b_1X_{i1} + b_2X_{i2} + \cdots + b_nX_{in}, \tag{1}$$

where P_i = the probability of bank i's failure, and $b = (b_1, \ldots, b_n)$ is a vector of regression coefficients for predictor variables X_i ($i = l, \ldots, n$). The logit model is preferred over the linear discriminant (MDA) model because it does not require multivariate normality among the independent variables and is computationally more tractable (see Espahbodi (1991, p. 56)). When the assumptions of MDA hold (namely, multivariate normality, equal variance-covariance matrices, and linearity), logit is equivalent to MDA.

Because of small sample sizes and the need to preserve degrees of freedom, we applied a stepwise logistic regression to select a subset of the most discriminating independent variables. Since no variables entered the model at the 10 percent significance level, we used a 30 percent significance level. The general lack of significance of the independent variables can be partially attributed to the small sample sizes (i.e., the cumulative distribution of the error terms in the regression relationship may not approximate a logistic function).

Logit models (and many other EWS methods) generate coefficient estimates for each of the variables and associated test statistics that indicate how well they discriminate between failed and non-failed banks. However, there are some potential drawbacks in terms of interpreting the results. For example, it is not possible to determine whether a significant variable is more useful in identifying failed banks than non-failed banks (i.e., no information about a variable's ability to reduce Type I versus Type II errors is available). Also, it is not possible to determine which of the variables is "out of line" for a particular bank. Such distinctions must be made in a univariate context by comparing mean values of variables in failed and non-failed banks.

Trait Recognition Approach

In this section, we briefly describe in general terms trait recognition as an EWS method.[6] *For a more in-depth description of the technique, see the appendix.* Building a trait recognition model is a multi-step process. The procedure involves: (1) selecting cutpoints for each of the variables, (2) assigning the variables binary codes, (3) constructing a trait matrix for each observation, (4) identifying good and bad traits (or distinctive features), and (5) selecting classification rules for the voting matrix. The model is evaluated for its ability to predict failed and non-failed banks in a holdout sample using the voting matrix. As already mentioned, trait recognition is a nonparametric EWS approach that relies on no statistical or distributional assumptions about the predictor variables.

A unique aspect of trait recognition is that it gathers and exploits information contained in complex interactions of variables. Individual *traits* are constructed from different segments of the distribution of each variable and the interactions of these segments with one or more other variables' segmented distributions. As an example, a trait of a failing bank could be a moderate level of return on assets and a high level of nonperforming loans as a proportion of total loans. Alternatively, a trait could be a moderate level of return on assets and a low level of equity capital to total assets. Notice that, upon dividing variables' distributions into low, middle, and upper segments, numerous interaction traits can be constructed between any two variables. This type of (binary) interaction variable captures different information than one can get by simply multiplying one variable by another variable, as typically done in finance, economics, and other fields of study.[7]

Once all possible traits of the variables are tabulated for all banks, trait recognition uses a search routine to cull traits that do not discriminate between failed and non-failed banks. Traditionally, EWS methods select one set of discriminatory variables; by contrast, trait recognition specifies two sets of discriminators: (1) safe traits associated with non-fail banks, and (2) unsafe traits associated with failing ones. These safe and unsafe traits are known as *features*. By tallying the number of safe and unsafe features for each bank, sample banks can be placed in a voting matrix defined by the number of safe and unsafe votes. Finally, the researcher selects rules for determining which cells in the voting matrix are safe (dominated by failed banks) or unsafe (dominated by non-failed banks) so that the observations can be used to classify and predict.

The trait recognition algorithm automatically builds the network of interactions with only a moderate amount of researcher input at various stages of its development (known as the *learning phase*). In contrast to neural networks in which interactions between variables are in a so-called hidden layer, trait recognition enables researchers to interpret the final results easily using a catalog of all interactions between variables in the model as reflected in the features. Unlike logit and other EWS methods, trait recognition produces variables (i.e., safe and unsafe features) that are clearly associated with either failed or non-failed banks. Simply by looking at the voting matrix, one can find a bank's position with respect to any variable as well as its standing relative to other banks. Since a record for each bank is provided that details its specific safe and unsafe features, it is obvious which features are out of line for any particular bank.

III. Empirical Methodology

Bank Samples

All U.S. commercial banks with assets greater than $250 million closed by the FDIC in the period 1989-1992 make up the large failed bank sample (n = 50). To create a matched-pair sample, we selected large non-failed banks for these failed banks according to the following criteria: (1) location in the same market as proxied by MSA and (2) nearest in asset size to the failed bank in the month of failure. The first criterion controls for size as a factor in failure. The second criterion controls for a difference in regulatory treatment and financial flexibility between large banks and small banks, as observed by Cole and Gunther (1994). Since our sample's total asset distribution ranges from $250 million to $14 billion, size differences can be substantial. While the matched-pair sampling design controls for such differences to some degree, we also conducted further analyses of the largest 30 failed banks (assets greater than $500 million), in addition to the largest 20 failed banks (assets greater than $700 million). These runs focus on progressively larger failed banks and further test the EWS capability of computer-based models using accounting data.

The second matching criterion attempts to control for different economic and competitive regional conditions among large banks across the country. During the period under study, most of the bank failures were concentrated in oil-dependent states (e.g., Texas) and areas where once-booming real estate had "busted" (e.g., the Northeast). Severe financial distress in some regional banking sectors no doubt disrupted the normal market behavior of many banks. These factors are discounted as much as possible by ensuring that each matched pair of failed and non-failed banks are from the same location.

Analyses are run both one year and two years prior to bank failures. While the failed bank sample remains the same in both of these runs, the matched-pair samples change because some paired non-failing banks do not pass the size-matching criterion. Out of 50 matched-pair non-failed banks in the one-year-prior sample, about 20 percent were replaced in the two-years-prior sample by another non-failing bank.

The matched pairs of 50 failed and 50 non-failed banks for one year and two years prior to failure were ordered by failed bank size and divided equally into an original sample and a holdout sample by using every other matched pair. The original samples were used to build the bank failure classification models, and the holdout samples were used to investigate the predictive power of the logit and trait recognition models.

Table 1 gives details of these bank samples. We have listed the names of the failed banks but not those of the non-failed banks (in order to avoid possible market reactions to our analyses of these banks). In most cases the matched non-failed bank's asset size is within 30 percent of that of the corresponding failed bank.

Independent Variables

Table 2 lists the independent variables and their mean values for the 25 failed and 25 non-failed banks in the original and holdout samples both one year prior and two years prior to the quarters in which failures occurred. Proxies for a variety of bank condition indicators are calculated from quarterly call report data, including size (X1, X2, and X28), profitability (X5 and X7), capitalization (X9), credit risk (X11, X13, X15, and X25), liquidity (X18), liabilities (X20 and X22), and diversification (X26). Size is measured in terms of both individual bank and holding company total assets, because the holding company is a source of strength that can reduce a bank's risk of failure. We also constructed a diversification measure by taking the sum of squared ratios of business loans, real estate, consumer loans, and securities to total assets.

These proxies and others in table 2 are fairly standard measures of bank condition that regulators, investors, and other interested parties monitor over time in performance evaluations.

In an effort to partially reflect the temporal behavior of the variables, variability measures were calculated. For example, based on data over the four quarters prior to failure for a particular bank, we calculated the maximum difference between quarters for a variable divided by its mean level over the four quarters. Presumably the financial performance of banks nearing collapse will vary from the norms. Since some banks have sudden changes in asset size prior to failure, we added another measure of size stability that used the mean change (rather than mean level) in total assets.

As shown in table 2, the mean values of the independent variables for the failed banks are significantly different from those for non-failed banks. Denoted by asterisks, the variables that have significant (at the 10 percent level) t-statistics for mean differences between failed and non-failed banks in all samples are X7 (profitability), X9 (capitalization), and X13 (credit risk). In addition, the variables X2, X5, X9, X15, X16, X20, X21, X22, and X25 are significantly different for failed and non-failed banks in at least one sample. Of these variables, X8, X10, X16, and X21 are variability measures and the remainder are in levels. Altogether, 14 out of 28 variables in table 2 have significantly different means in one or more of the four sample sets of data.

Parametric models such as logit tend to select variables based on the strength of the statistical relationship characterized in univariate tests. The 14 variables cited above are most likely to be prominent variables in the logit models. By contrast, nonparametric models such as trait recognition may well be expected to be influenced by other predictors because they are not grounded in statistical properties.

IV. Empirical Results

Logit Models

Table 3 reports the estimated logit model for each bank sample and related classification and prediction results. The models employed about 40 percent of the 28 independent variables. As expected, the variables with significant univariate t-statistics for mean differences in table 2 are frequently significant in the multivariate logit models.

As shown in table 3, the classification results for the logit models using the original samples are quite strong using data from one year prior to failure (panel A) and two years prior (panel B). Based on the original sample (i.e., n = 50) and smaller samples of progressively larger banks (i.e., n = 30 and n = 20), four out of six logit models obtained 100 percent correct classification. The remaining logit models obtained at least 90 percent correct classification. These results suggest that the logit approach performs well within sample.

Table 3 also shows that the logit prediction results using the holdout samples are moderately accurate, with the percentage correct falling to between 60 percent and 75 percent. In the n = 50 and n = 30 samples the one-year-prior models outperformed the two-years-prior models, and vice versa for the n = 20 sample. Since the results for one-year-prior and two-years-prior models are mixed, we infer that the model's predictive ability was not substantially affected by extending the forecasting horizon from one year to two years. Unfortunately, the EWS efficacy of the logit models is suspect because between 40 percent and 65 percent of the failed banks are not predicted to fail (Type I errors). Because of the magnitude of such errors, one must question the value of the model to bank regulatory agencies. We infer that, although the logit models have a moderate degree of overall predictive power, their out-of-sample performance is not better than random chance in the case of large bank failures.

Trait Recognition Models

Tables 4 to 6 contain the results for the trait recognition models. For purposes of documentation, table 4 lists the trait recognition cutpoints for each of the 28 independent variables using the original sample (n = 50). Cutpoints were automatically generated one standard deviation above and below the variable means. As discussed in the previous section, two cutpoints for each variable divide their distributions into three segments for purposes of binary coding of the variables and subsequent construction of binary strings for each bank.

Table 5 provides further documentation of the trait recognition models by listing the safe and unsafe features employed in the one-year-prior and two-years-prior models (as shown in panels A and B, respectively) for the original sample (n = 50). Notice that the only variable that is not incorporated in the features is X28, which is calculated as the ratio X1/X2 (or bank total assets divided by holding company total assets). Unlike the logit models, the trait recognition method employed all variables as predictors, with the exception of variables with redundant information contained in interactions of other variables.

Table 5 shows that more than 50 percent of the safe and unsafe features contain three-variable interactions (i.e., 44 out of 71 features and 26 out of 46 features in the one-year-prior and two-years-prior models, respectively). Most other features are two-variable interactions, with few features defined by a single variable. These results suggest that most valuable information in identifying large failing banks is contained in complex variable interactions.[8] However, most EMS methods, with the exception of artificial neural networks with multiple interactive layers, do not exploit the information contained in variable interactions. Even if these methods are run with a comprehensive list of interaction variables, as already discussed, they still would not parallel the complex interactions used in the traits model.[9]

Table 6 reports the classification and prediction results for the trait recognition models. All six original sample runs (i.e., one-year-prior and two-years-prior data models using n = 50, n = 30, and n = 20 banks) obtained 100 percent correct classification. Concerning predictive ability on the holdout samples, the trait recognition models correctly identified from 63 percent to 95 percent of large banks. In contrast to the logit results, the one-year-prior results in panel A of table 6 (i.e., 80 percent to 95 percent predictive accuracy) consistently attained a higher predictive accuracy than the two-years-prior results in panel B (i.e., 63 percent to 70 percent). Also unlike the logit models' results, the trait recognition models tended to perform well above chance in predicting failing banks. In the one-year-prior models 23 out of 25, 11 out of 15 and 10 out of 10 failures are correctly predicted in the samples with n = 50, n = 30, and n = 20, respectively. In the two-years-prior models, 17 out of 25, 10 out of 15, and 8 out of 10 are correctly predicted in the aforementioned sample sizes, respectively. These Type I error results are considerably stronger than those obtained from the logit models, which had considerable difficulty in predicting large bank failures in the holdout samples. Notably, in all six test cases using one-year-prior and two-years-prior data and different sample sizes, trait recognition correctly predicted a greater number of failures than logit.

Weighted Efficiency Scores

In table 7 we report the weighted efficiency (WE) scores for logit versus trait recognition models. The WE scores, as noted by Espahbodie (1991), adjust overall prediction rates for the fact that Type I errors are more serious than Type II errors. The WE scores indicate that the trait recognition models outperformed the logit models in all samples. The former models did especially well using one-year-prior data. (For example, based on the n = 20 sample containing the largest banks, the one-year-prior WE score is 86.4 for the trait recognition model and 24.0 for the logit model.) These results suggest that the trait recognition model may be an effective early warning system (EWS) for large banks.

Stability of Predictions for Individual Large Banks

Lastly, details of the prediction results for each of the failed banks in the holdout samples based on the logit and trait recognition models are presented in tables 8 and 9, respectively. In table 8 we report the predicted probabilities of failure estimated by the logit models using the different sample sizes. The incorrectly classified failing banks are marked with an asterisk. Given that the banks are shown in rank order, it is apparent that there is no systematic bias toward missing the largest bank failures in the samples. However, an unfavorable pattern in the predicted probabilities for the failed banks is that they are normally either 1.0 or 0.0 (i.e., no chance of survival or failure, respectively). Because estimated probabilities of failure are consistently extreme the logit results may not be very reliable. Also, the logit models give inconsistent predictions for individual large failed banks using one-year-prior and two-prior models – that is, both logit models correctly predicted only seven failures in the n = 25 failed bank sample, four failures in the n = 15 failed bank sample, and three failures in the n = 10 failed bank sample. Thus, not only do the one-year-prior and two-years-prior models perform no better than chance in predicting large bank failures, the predictions are not stable for any particular large bank as the failure event approaches.

Table 9 details the number of safe and unsafe votes obtained for each large failed bank using the trait recognition models for the different sample sizes. As before, missed predictions are marked with an asterisk. While some of the misclassified failed banks had no unsafe votes (i.e., tantamount to zero probability of failure using logit), most of the misclassified banks did have one or more unsafe votes. Also, for the correctly identified failed banks, the number of safe and unsafe votes varies considerably from one sample to another, and these differences imply greater or lesser risk of failure. Hence, unlike the logit results, the trait recognition results can convey the degree to which the risk of failure is growing. Finally, comparing the predictions based on the one-year-prior and two-years-prior models, in the n = 25, n = 15, and n =10 samples of failed banks, the number of failures predicted by both models is 18, 8, and 8, respectively. These results are more stable than those for logit and suggest that the trait recognition models give fairly consistent EWS signals about large banks using one-year-prior and two-years-prior data.

IV. Summary and Conclusions

The present study empirically examined the efficacy of using computer-based EWS models to assess the risk of failure of large U.S. commercial banks. Because of a lack of sufficient data points in the past, few studies have been published on this subject, with the exception of studies of stock prices and failure risk and case studies of individual large bank failures. In the late 1980s and early 1990s, numerous large banks failed in the United States. These failures allowed us to gather a sample of 50 large banks with more than $250 million in assets. To them, we added a matched-pair sample of 50 large non-failed banks. Using these data, we compared the predictive ability of logit analysis, a parametric approach, with that of trait recognition, a nonparametric approach.

While both EWS methods performed well in terms of classification accuracy using original samples, trait recognition outperformed logit in a variety of tests using holdout samples. Based on one-year-prior and two-years-prior models and different holdout samples, the predictive accuracy of the logit models was moderately successful. By comparison, the predictions of trait recognition models were substantially accurate using one-year-prior and two-years-prior data. Trait recognition was able to predict most of the failures both one and two years prior to collapse. And trait recognition's ability to identify a higher number of failing large banks than logit was not achieved by misclassifying a greater number of non-failing banks. The weighted efficiency scores for trait recognition models surpassed those of logit models in all test cases. Further results from the evaluation of the individual large banks' risks of failure one year and two years prior to failure suggested that trait recognition provided more stable EWS signals than logit in the years preceding failure.

We conclude that trait recognition is a potentially useful early warning system for large failing banks. Apparently, complex two- and three-variable interactions captured by the trait recognition method contain valuable information about large bank failure risk. In light of the ongoing consolidation in the banking industry in the United States and other countries, further studies should investigate large banks' risks of failure. Research on predicting capital inadequacy (as defined by regulatory standards), securities investment losses, and other types of bank risk among large banks would be a logical extension of the bank EWS literature.

We believe EWS models developed on a one-year-prior or two-years-prior performance window are useful as supervisory tools. Accurately identifying large banks that are likely to fail within the next two years may provide sufficient time for supervisors to impose restrictions that reduce resolution costs. However, a one- or two-year window may be too narrow for implementing corrective actions to avert failure/closure. Models that identify problem banks three years or even four years prior to failure/closure may be more useful as EWS tools for the purpose of implementing corrective actions to avoid failure. However, identifying problem banks so early in their decline using only publicly available data is much more difficult empirically. Selecting reliable samples and designing a practicable model will be formidable tasks.

Footnotes

1. See Altman and Saunders (1998) for an excellent overview of this literature and additional citations.

2. Bank failure or closure is a supervisory action that occurs when the regulators recognize book-valued insolvency, a condition that may not be consistent with the timing of market-valued insolvency. It is essentially an administrative option that reflects the condition of the bank but is applied at the discretion of the banking authorities (see Demirguc-Kunt (1992) for a more detailed discussion on the differences between insolvency, failure, and closure). The failure model presented below does not predict either book- or market-valued insolvency per se. That is, the model is not designed to measure or forecast the solvency of specific institutions or attempt to identify factors that suggest a bank will be insolvent in the near future. A more complex analysis of the expected market value of equity would be required to address that issue. The development of an empirical insolvency-based prediction model is made more difficult because the regulator may intervene to force changes that may avert insolvency. Also, the intangible value of a bank's liabilities is difficult to measure from industry-level data such as call report data but can be important when troubled banks are purchased.

3. Another possible policy response to large bank fragility and the potential for systemic risk is to increase market discipline (see Kaufman (1999)).

4. According to FDIC *Annual Reports*, the total number of U.S. failed banks (by year) in this period was as follows: 207 (1989), 169 (1990), 127 (1991), and 127 (1992). This pace of failures far surpassed the historical average of fewer than 25 banks per year.

5. This small sample facility of trait recognition is also apparent in previous applications to geophysics problems involving the prediction of seismic risk (e.g., see Gelfand *et al.* (1976), Briggs, Press, and Guberman (1977), Caputo *et al*. (1980), Benavidez and Caputo (1988)) and mineral deposits (e.g., see Bongard *et al*. (1966) and Briggs and Press (1977)).

6. In the present study we employ the Bongard *et al.* (1966) algorithm used in geophysical studies, which has been generalized by Kolari *et al*. (1966) to finance and economics problems.

7. As a further example, the traits method explores numerous interactions between any two variables X1 and X2. Given the use of two cutpoints for each variable's distribution, the interactions between the segments of the two variables' distributions can be written as follows (where L = lower segment of the variable distribution, LM = lower and middle segments, MU = middle and upper segments, and U = upper segment):

X1L/X2L	X1LM/X2L	X1MU/X2L	X1U/X2L
X1L/X2LM	X1LM/X2LM	X1MU/X2LM	X1U/X2LM
X1L/X2MU	X1LM/X2MU	X1MU/X2MU	X1U/X2MU
X1L/X2U	X1LM/X2U	X1MU/X2U	X1U/X2U

If three variables are considered, it is obvious that a large number of complex interactions are possible between the different segments of their distributions.

8. By inference, it seems plausible that a similar finding is possible in other areas of finance that use accounting and financial data to predict such outcomes as private nonfinancial firm failure, bond ratings, etc.

9. Given the success of the interaction variables in the trait recognition models, we contemplated including interaction variables in the logit models. However, while the 28 predictors can be taken two at a time and then three at a time to define a large number of interaction variables (more than 1,000 new variable definitions), this type of exhaustive exploratory search is not consistent with its application in the finance and banking literature. Nor is it designed to explore a vast number of possible variable interactions. Also, as mentioned, the standard practice of constructing interaction variables by means of multiplicative transformations would not capture the more complex and numerous interactions captured by trait recognition.

References

Altman, E. (1968), “Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *Journal of Finance* 23, pp. 589-609.

Altman, E, R. Haldeman, and P. Narayanan, (1977), “ZETA Analysis: A New Model to Identify Bankruptcy Risk of Corporations,” *Journal of Banking and Finance* 1, pp. 29-54.

Altman, E. (1993), *Corporate Financial Distress and Bankruptcy*, 2nd edition, John Wiley & Sons: New York, NY.

Altman, E. I., G. Marco, and F. Varetto (1994), “Corporate Distress Diagnosis: Comparisons Using Linear Discriminant Analysis and Neural Networks (the Italian Experience),” *Journal of Banking and Finance* 18, pp. 505-529.

Altman, E. I. and P. Narayanan (1997), “Business Failure Classification Models: An International Survey,” in Choi, F. (Ed.), *International Accounting*, 2nd edition, New York: Wiley.

Altman, E. I. and A. Saunders (1998), “Credit Risk Measurement: Developments Over the Last 20 Years,” *Journal of Banking and Finance* 21, pp. 1721-1742.

Beaver, W.H. (1966), “Financial Ratios as Predictors of Failure,” in *Empirical Research in Accounting: Selected Studies, 1966; Journal of Accounting Research* 4 (Supplement), pp. 71-111.

Beaver, W.H. (1968), “Market Prices, Financial Ratios, and the Prediction of Failure,” *Journal of Accounting Research* 6, pp. 179-192.

Benavidez, A., and M. Caputo (1988), “Pattern Recognition of Earthquake Prone Areas in the Andes,” *Revisita Geofisica*, pp. 141-163.

Berger, A. N., A. K. Kashyap, and J. M. Scalise (1995), "The Transformation of the U.S. Banking Industry: What a Long, Strange Trip It's Been," *Brookings Papers on Economic Activity* 2.

Blum, M. (1974), “Failing Company Discriminant Analysis,” *Journal of Accounting Research* 12, pp. 1-25.

Bongard, M. M., M. I. Vaintsveig, S. A. Guberman, and M. L. Izvekova (1966), “The Use of Self Learning Programs in the Detection of Oil Containing Layers,” *Geology Geofiz* 6, pp. 96-105.

Bovenzi, J. F., J. A. Marino, and F. E. McFadden (1983), “Commercial Bank Failure Prediction Models,” *Economic Review*, Federal Reserve Bank of Atlanta, pp. 186-195.

References, continued

Boyd, J. H. and S. L. Graham (1996), "Consolidation in U.S. Banking: Implications for Efficiency and Risk," Working Paper No. 572, Federal Reserve Bank of Minneapolis.

Breiman, L., J. H. Freidman, R. A. Olshen, and C. J. Stone (1984), *Classification and Regression Trees* , Wadsworth International Group: Belmont, CA.

Briggs, P., and F. Press (1977), "Pattern Recognition Applied to Uranium Prospecting," *Nature* 268, pp. 125-127.

Briggs, P., F. Press, and S. A. Guberman (1977), "Pattern Recognition Applied to Earthquake Epicenters in California and Nevada," *Geological Society of America Bulletin* 88, pp. 161-173.

Caputo, M., V. Keilis Borok, E. Oficerova, E. Ranzman, I. I. Rotwain, and A. Solovieff (1980), *Physics Earth and Planetary Interiors*, 21, pp. 305-320.

Coats, P. and L. Fant (1993), "Recognizing Financial Distress Patterns Using a Neural Network Tool," *Financial Management* 22, pp. 142-155.

Cole, R. A. and J. W. Gunther (1994), "When Are Failing Banks Closed?" *Financial Industry Studies*, Federal Reserve Bank of Dallas (December), pp. 1-12.

Demirguc-Kunt, A. (1992), "Deposit-Institution Failures: A Review of Empirical Literature," *Economic Review*, Federal Reserve Bank of Cleveland, pp. 2-18.

Edmister, R. O. (1972), "An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction," *Journal of Financial and Quantitative Analysis* 7, pp. 1477-1494.

Eisenbeis, R. A., (1977), "Pitfalls in the Application of Discriminant Analysis in Business, Finance, and Economics," *The Journal of Finance* 32, pp. 875-898.

Espahbodi, P. (1991), "Identification of Problem Banks and Binary Choice Models," *Journal of Banking and Finance* 15, pp. 53-71.

Federal Deposit Insurance Corporation (1997), *History of the Eighties – Lessons for the Future*, Vol. 1, Washington, D.C.

Frydman, H., E. I. Altman, and D. Kao (1985), "Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress," *The Journal of Finance* 40, pp. 269-291.

Gajewski. G. R. (1989), "Assessing the Risk of Bank Failure," *Proceedings from a Conference on Bank Structure and Competition*, Federal Reserve Bank of Chicago, pp. 432-456.

References, continued

Gelfand, I. M., Sh. A. Guberman, M. L. Izvekova, V. I. Keilis Borok, L. Knopoff, E. Ranzman, F. Press, E. Ranzman, M. Rotwain, and A. M. Sadowsky (1976), “Criteria of High Seismicity Determined by Pattern Recognition,” In Ritzema (ed.). The Upper Mantle, *Tectonophysics* 13, pp. 415-422.

Hoenig, T. M. (1999), “Financial Industry Megamergers and Policy Changes,” *Economic Review* 84, Federal Reserve Bank of Kansas City, pp. 7-14.

Jones, F. (1987), “Current Techniques in Bankruptcy Prediction,” *Journal of Accounting Literature* 6, pp. 131-164.

Kaufman, George G. (1999), “Banking and Currency Crises and Systemic Risk: A Taxonomy and Review,” Working Paper No. WP-99-12, Federal Reserve Bank of Chicago.

Keasey, K. and P. McGuinnes (1990), “The Failure of UK Industrial Firms for the Period 1976-1984, Logistic Analysis and Entropy Measures,” *Journal of Business, Finance and Accounting* 17, 119-135.

Kolari, J., M. Caputo, and D. Wagner (1996), “Trait Recognition: An Alternative Approach to Early Warning Systems in Commercial Banking,” *Journal of Business, Finance & Accounting* 23, pp 1415-1434.

Korobrow, L. and D. Stuhr (1985), “Performance Measurement of Early Warning Models: Comments on West and Other Weakness/Failure Prediction Models,” *Journal of Banking and Finance*, June, pp 267-273.

Lane, W. R., S. W. Looney, and J. W. Wansley (1986), “An Application of the Cox Proportional Hazards Model to Bank Failure,” *Journal of Banking and Finance* 10, pp. 511-531.

Maddala, G. (1986), “Econometric Issues in the Empirical Analysis of Thrift Institutions’ Insolvency and Failure,” Working Paper No. 56, Federal Home Loan Bank Board.

Martin, D. (1977), “Early Warning of Bank Failure: A Logit Regression Approach,” *Journal of Banking and Finance* 1, pp. 249-276.

Meyer, P. and H. Pifer (1970), “Prediction of Bank Failures,” *The Journal of Finance* 25, pp. 853-868.

Ohlson, J. (1980), “Financial Ratios and the Probabilistic Prediction of Bankruptcy,” *Journal of Accounting Research* 18, pp. 109-131.

References, continued

Peavy, J. W., III and G. H. Hempel (1988), "The Penn Square Bank Failure: Effect on Commercial Bank Security Returns – A Note," *Journal of Banking and Finance* 12, pp. 141-150.

Pettway, R. H. (1976), "Market Tests of Capital Adequacy of Large Commercial Banks," *The Journal of Finance* 31, pp. 865-875.

Pettway, R. H. (1980), "Potential Insolvency, Market Efficiency, and Bank Regulation of Large Commercial Banks," *Journal of Financial and Quantitative Analysis* 15, pp. 219-236.

Platt, H. D. and M. B. Platt (1990), "Development of a Class of Stable Predictive Variables: The Case of Bankruptcy Prediction," *Journal of Business, Finance and Accounting* 17, pp. 31-51.

Santomero, A. and J. Vinso (1977), "Estimating the Probability of Failure for Firms in the Banking System," *Journal of Banking and Finance* 1, pp. 185-206.

SAS User's Guide: Statistics (1985) SAS Institute: Cary, North Carolina.

Sinkey, J. E., Jr. (1975), "A Multivariate Statistical Analysis of the Characteristics of Problem Banks," *Journal of Finance* 20, pp. 21-36.

Sinkey, J. E., Jr. (1985), "The Characteristics of Large Problem and Failed Banks," *Issues in Bank Regulation* (Winter), pp. 43-53.

Stuhr, D. P. and R. Van Wicklin (1974), "Rating the Financial Condition of Banks: A Statistical Approach to Aid Bank Supervision," *Monthly Review*, Federal Reserve Bank of New York, September, pp. 233-238.

Thomson, J. B. (1991), "Predicting Bank Failures in the 1980s," *Economic Review* 27, Federal Reserve Bank of Cleveland, pp. 9-20.

West, R. C. (1985), "A Factor-Analytic Approach to Bank Condition," *Journal of Banking and Finance* 9, pp. 253-266.

Whalen, G. and J. B. Thomson (1988), "Using Financial Data to Identify Changes in Bank Condition," *Economic Review* 24, Federal Reserve Bank of Cleveland, pp. 17-26.

Zavgren, C. (1985), "Assessing the Vulnerability of Failure of American Industrial Firms: A Logistic Analysis," *Journal of Business, Finance and Accounting* 12, pp. 19-45.

Appendix. Description of Trait Recognition

In this section we describe the steps involved in building a trait recognition model (see Kolari *et al*. (1996)). Figures 1 to 3 give a simple example for illustrative purposes. We assume that the sample consists of five failed banks (denoted a, b, c, d, and e) and five non-failed banks (denoted A, B, C, D, and E). Also, we assume that three financial ratios – namely, X1 = net income/total assets, X2 = loan losses/total assets, and X3 = equity capital/total assets – are used to measure bank financial condition. Each ratio is calculated one year prior to the collapse of the failed banks for the 10 banks under examination.

Step 1 – Recoding the variables into binary strings. Recoding can be done using a general partitioning rule or researcher intuition. A general partitioning rule may take the following form: one standard deviation around the sample mean is calculated and the distribution of values for a particular variable is partitioned into three segments using two *cutpoints* at plus or minus one standard deviation. The lower, middle, and upper segments of the distribution defined by the two cutpoints are coded, 00, 01, 11, respectively. Given j = 1, …, n banks and i = 1, …., p variables, the jth bank's vector of variables $X_{ij} = [X_1, \ldots, X_p]$ is recoded into binary form $X_{ij} = [B_1, B_2, \ldots. B_L]$, where L is the length of the string and two digits describe each variable. For example, variable vector [X1, X2, X3] = [010011] implies that bank j is in the middle (01), low (00), and upper (11) segments of the distributions of X1, X2, and X3, respectively.

Alternatively, the researcher could produce graphs of the distributions of the variables and then visually select two cutpoints for each variable. In this way the researcher derives the cutpoints based on experience and intuition. This option provides the researcher with a good grasp of the raw data entering the model, as well as control over the binary coding process. Researcher judgment is used to set thresholds at which the value of a variable is considered low or high (e.g., regulatory practice applies specific capital ratios to divide banks into well-capitalized, adequately capitalized, and undercapitalized groups). The general notion is to set the cutpoints such that the low segment has primarily (for example) failed banks and vice versa for the upper segment. The middle segment is a mix of failed and non-failed banks. At times only one cutpoint is possible because there is no clear middle segment (i.e., only low and upper segments are coded). Also, some variables have distributions that have failed and non-failed banks mixed throughout their distribution. In those cases the variable can either be dropped or a cutpoint can be selected based on theory or practice. We illustrate the recoding process in panel A of figure 1. In this example the distributions for X1, X2, and X3, as well as the cutpoints selected by the researcher, are presented. Based on figure 1, the 10 sample banks are coded into binary strings, as shown in panel B. Finally, it should be noted that more than three segments are possible and a different binary code could be assigned to each region (e.g., a three-digit code); however, previous research has suggested that having more than three segments does not substantively increase predictive power.

While the binary coding of a continuous variable would seem to throw out cardinal level data, the cost of this loss of information is offset by: (1) the ability to construct a variety of complex interaction variables that are not possible with cardinal data (to be discussed shortly), and (2) the advantage of applying fuzzy logic to capture general patterns in the data. Regarding the latter advantage, because the model is dependent on general patterns in the variables, trait recognition is less data-sensitive than models using cardinal variable measurements to outlier biases. This emphasis on general patterns may contribute to a more stable model over time than methods that focus on small increments in variables.

Step 2 – The binary strings are converted to trait matrices. Notice that the individual bank binary strings in the trait matrix have different *patterns*, with the exception of failed banks c and d with the identical string 011100. Furthermore, there appear to be patterns in these binary strings that distinguish failed from non-failed banks; for example, X1 (or profit rate) and X3 (or capital ratio) tend to have 0 codes for failed banks and 1 codes for non-failed banks, whereas X2 (or loans/assets ratio) normally has 1 codes for failed banks and 0 codes for non-failed banks. These different patterns suggest that the binary strings may be useful in discriminating between failed and non-failed banks.

Figure 2 gives details of how binary strings use fuzzy logic to define the patterns among observations under study. The value of a digit in the binary string gives the general location of the observation (bank) in each variable's distribution. For example, if the first digit of variable X1 equals 1, this implies that the bank's rate of return on assets is in the upper segment of the distribution, as determined by the cutpoints. Alternatively, a 1 for the second digit of X1 coincides with either the middle or upper segments of the distribution. We denote the five possible locations in any variable's distribution as follows: lower (L), middle (M), upper (U), lower/middle (LM), and middle/upper (MU). Figure 2 gives an example of how to read the data encoded in a binary string for failed bank e, as defined in panel B of figure 1.

The string of binary codes is recoded to explore interactions between patterns within the binary strings. To do this a matrix of traits for each bank is created from its binary string. The traits matrix considers all possible combinations of the variables taken one, two, and three at a time and, therefore, catalogues all interrelations between the variables. More specifically, each trait (T) contains an array of six integers, T = p,q,r,P,Q,R, where p, q, and r are *pointers* to positions in the binary string, and P, Q, and R give the *values* of the binary code at their corresponding positions p, q, and r. The basic rules for constructing the traits matrix for binary string of length L are as follows: p = 1, …, L; q = p, p+1, …, L; r = q, q+1, …, L; P = 0 or 1; Q = 0 or 1; and R = 0 or 1.

For example, the trait matrix for failed bank e with binary string 110100 can be created by considering all possible combinations of the six digits taken one, two, and three at a time as follows:

p q r PQR	p q r PQR	p q r PQR
1 1 1 111	2 3 3 100	1 2 4 111
2 2 2 111	2 4 4 111	1 2 5 110
3 3 3 000	2 5 5 100	1 2 6 110
4 4 4 111	2 6 6 100	2 3 4 101
5 5 5 000	3 4 4 011	2 3 5 100
6 6 6 000	3 5 5 000	2 3 6 100
1 2 2 111	3 6 6 000	2 4 5 110
1 3 3 100	4 5 5 100	2 4 6 110
1 4 4 111	4 6 6 100	3 4 5 010
1 5 5 100	5 6 6 000	3 4 6 010
1 6 6 100	1 2 3 110	4 5 6 100

As shown above, there are 33 traits for this six-digit string.[4] Notice that the traits matrix allows for all possible interactions of the segmented variables' distributions. The trait 111111 points only to variable X1, where the location of the bank is the upper segment of X1's distribution. Trait 133100 points to variables X1 and X2, where X1 is in the upper segment of its distribution, and X2 is in the lower or middle segments of its distribution (see figure 2). Likewise, trait 236100 points to variables X1, X2, and X3, where X1 is in the middle or upper segments, X2 is in the lower or middle segments, and X3 is in the lower segment. Notice that the combination 236100 is identical to 632001, such that the latter trait is redundant and can be dropped from consideration. A different traits matrix is generated for each bank with a different binary string.

Step 3 – The features in the trait matrices are retained. Since the size of the traits matrix increases greatly as the number of digits in the binary string increases (because of factorial mathematics), and each bank has its own traits matrix, even problems with only 10 independent variables can quickly exhaust computer disk space. So, only traits that are useful in discriminating between failing and non-failed banks are retained. These traits are known as *features*. A *safe feature* is a trait that is present frequently in non-failed banks but infrequently in failed banks, and vice versa for *unsafe features.* For example, if a particular trait (or six-digit sequence of binary codes) is found in a minimum of 75 percent of the failed banks and a maximum of 25 percent of the non-failed banks, it could be designated as an unsafe feature. (The percentages defining safe features and unsafe ones are of necessity ad hoc.) Previous work by Kolari *et al*., as well as our own experience, suggests that at least five safe and five unsafe features (i.e., 10 total features) are needed to discriminate between failed and non-failed banks. Also, using as many as 100 safe or unsafe features is excessive and does not make the model more discriminating. In general, the researcher should experiment with relaxing minimum and maximum percentage limits to increase the number of safe and unsafe features and observe whether the model is more discriminating. Setting strict percentage limits tends to throw out valuable information that can increase the model's performance, while lenient limits collects excess information with marginal incremental discriminatory power (or no added predictive value). Finally, the traits program filters out features that duplicate the observations of some other feature.

Step 4 – The safe and unsafe features are used to vote on each bank. At this point the remaining safe and unsafe features can be used to "vote" on each bank in the sample and then classify the banks. The numbers of safe and unsafe votes for each bank are tallied and a voting matrix is constructed. Panel A of figure 3 gives a hypothetical list of three unsafe and three safe features. Panel B shows the voting results for the sample banks defined in figure 1. Finally, panel C illustrates the voting matrix for the sample banks. If the failed banks outnumber the non-failed banks in a cell, the cell is unsafe; if the failed banks outnumber the non-failed ones, the cell is safe. If no banks fall in a cell, it is not classified. For cells with equal numbers of failed and non-failed banks and for unclassified cells, if the safe votes outnumber the unsafe, the cell is safe; if unsafe votes outnumber the safe, the cell is unsafe. Other cells can be manually classified as safe or unsafe according to researcher judgement. If no banks fall in mixed cells, a 100 percent correct classification rate will be obtained. Only if banks are located in mixed cells will the overall classification rate be less than 100 percent. By running the traits program a few times with different minimum and maximum percentage limits in the features selection step, a voting matrix with no mixed cells can normally be achieved. Thus, unlike virtually all other EWSs, the original sample traits program typically yields 100 percent correct classification.

If there are few observations and the number of features is quite high so that a large voting matrix is created (e.g., 60x60 matrix with 3,600 cells but only 40 observations), most cells will be unclassified and the aforementioned simple decision rule will be applied to most observations in a holdout sample. The holdout sample classification results are weaker in this instance than could be obtained by using a smaller voting matrix with fewer unclassified cells and a greater proportion of cells classified as safe or unsafe based on the original sample data. In effect, concrete decision rules determined by the original sample data are more effective in holdout sample prediction than naïve rules that simply count the number of safe and unsafe votes and then classify a cell on that basis alone. In this regard, a benefit of the voting matrix is a visualization of the pattern of failed and non-failed banks in the two-dimensional space defined by safe and unsafe features. One might intuit from panel C of figure 3, by casual inspection it is intuitively obvious that banks with two or more unsafe features are very likely failing and that banks with one unsafe feature are suspect. The voting matrix enables the analyst to quickly understand where an individual bank under investigation lies in the safe and unsafe features space relative to other banks.

Step 5 – A two-stage model is developed under certain conditions. We added another step to the traits program analysis if a considerable proportion of our sample banks inhabited mixed cells. In Kolari *et al*. (1996) the classification results were improved by using multiple stages of traits models. The first stage model is used to classify banks that are easily identified as failing or non-failing. Banks that fall in mixed cells (or gray areas) are held out as a separate subsample, and an entirely new traits model is developed to identify these banks. This procedure takes advantage of the fact that the traits program works well with small numbers of observations. Additionally, the program focuses attention on banks in gray areas, which are the crux of most classification problems. We arbitrarily set the minimum number of banks in mixed cells at 10 failed banks and 10 non-failed banks to allow sufficient observations to build a second model if needed.

Of course, any model developed from limited numbers of observations is narrowly specified and may well yield incorrect predictions. One way to mitigate this shortfall is to continually update the model for any unpredicted failures (i.e., Type I errors) by tagging the cell in which the failed bank falls, such that a repeat of the same mistake is avoided. This learning process of updating and improving the traits model is intended to continually strengthen the model as it gains further experience with the objects under study.

Figure 1
Selection of Cutpoints Based on Distributions of
Financial Ratios and Subsequent Binary Strings Coding of Banks

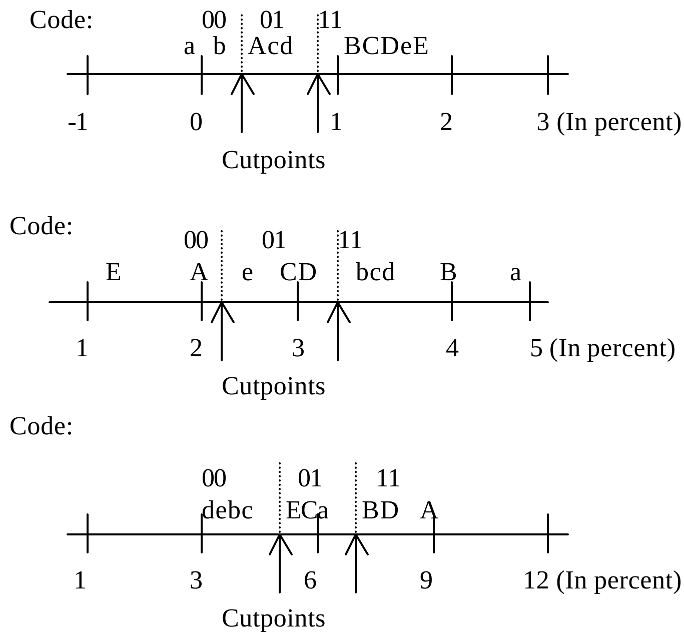
Given the following sample banks: Failed banks: a, b, c, d, e
Non-failed banks: A, B, C, D, E

A. Selection of cutpoints in distributions of variables

$X_1 = \frac{\text{net income}}{\text{total assets}}$

$X_2 = \frac{\text{loan assets}}{\text{total assets}}$

$X_3 = \frac{\text{equity capital}}{\text{total assets}}$



B. Binary strings coding of observations

Failed banks

	X_1		X_2		X_3	
a	0	0	1	1	0	1
b	0	0	1	1	0	0
c	0	1	1	1	0	0
d	0	1	1	1	0	0
e	1	1	0	1	0	0

Non-failed banks

	X_1		X_2		X_3	
A	0	1	0	0	1	1
B	1	1	1	1	1	1
C	1	1	0	1	0	1
D	1	1	0	1	1	1
E	1	1	0	0	0	1

Figure 2
Fuzzy Logic of Binary Strings

A. Binary strings and variables (X_j)

Binary string: 110100
Positions (pqr): 123456
X_1 = positions 1 and 2
X_2 = positions 3 and 4
X_3 = positions 5 and 6

B. Binary strings and segments of variable distributions

For each variable X_1, X_2, and X_3, and given the following coding scheme based on the distribution of observations as described in figure 1:

00 (Lower) | 01 (Middle) | 11 (Upper)

- If the first digit = 0, the observation is in the lower or middle (LM) segment of the variable distribution;
- If the first digit = 1, the observation is in the upper (U) segment of the variable distribution;
- If the second digit = 0, the observation is in the lower (L) segment of the variable distribution; and
- If the second digit = 1, the observation is in the middle or upper (MU) segment of the variable distribution.

C. Interpreting binary strings

For the binary string of the individual bank given above in panel A:

Positions (pqr)	1	2	3	4	5	6
Binary string values (PQR)	1	1	0	1	0	0
Variables	X_1	X_1	X_2	X_2	X_3	X_3
Segment of variable distribution	U	MU	LM	MU	LM	L

Figure 3

Using Features to Construct the Voting Matrix

A. Interpreting features

Features:

No.	Type	p	q	r	P	Q	R
#1.	Unsafe	2	5	6	0	0	0
#2.	Unsafe	2	3	5	0	1	0
#3.	Unsafe	1	2	6	0	1	0
#4.	Safe	1	2	5	1	1	1
#5.	Safe	2	2	5	1	1	0
#6.	Safe	2	3	6	1	0	1

Relationship of the features to:

Variables			Segments		
X_1	X_3	X_3	L	LM	L
X_1	X_2	X_3	L	MU	LM
X_1	X_1	X_3	LM	L	L
X_1	X_1	X_3	U	MU	U
X_1	X_1	X_3	MU	MU	LM
X_1	X_2	X_2	MU	LM	MU

L = Lower segment
M = Middle segment
U = Upper segment

LM = Lower or middle segment
MU = Middle or upper segment

B. Voting on binary strings of banks based on features

Failed banks							Votes: Safe	Votes: Unsafe	Features
a	0	0	1	1	0	1	0	1	#2
b	0	0	1	1	0	0	0	2	#1,#2
c	0	1	1	1	0	0	1	1	#3,#5
d	0	0	1	1	0	0	0	2	#1,#2
e	1	1	0	1	0	0	1	0	#5

Non-failed banks							Votes: Safe	Votes: Unsafe	Features
A	0	1	0	0	1	1	1	0	#6
B	1	1	1	1	1	1	1	0	#4
C	1	1	0	1	0	1	2	0	#5,#6
D	1	1	0	1	1	1	2	0	#4,#6
E	1	1	0	0	0	1	2	0	#4,#6

C. Voting matrix

Unsafe Votes

Safe Votes	0	1	2	3
0		a	b, d	
1	e, A, B	c		
2	C, D, E			
3				

Safe Cells: (2, 0)
Unsafe Cells: (0, 1), (0, 2), (1, 1)
Mixed Cells: (1, 0)
Unclassified Cells: (0, 0), (0, 3), (1, 2), (1,3), (2, 1), (2, 2), (2, 3), (3, 0), (3, 1), (3, 2), (3, 3)

Table 1. Large Failed Banks in Original and Holdout Samples

A. Original Sample (n = 25)

Bank ID	Name	MSA	Bank Assets[a]	BHC Assets[a]	Failure Date
652034	Southeast Bank NA	5000	15469836	15469836	09/91
870351	Mbank, Dallas NA	1920	7309953	22339037	03/89
653657	MBank Houston NA	3360	4828446	22339037	03/89
318554	Texas AMN Bank, Fort Worth NA	2800	2667164	5817439	09/89
784225	National Bank of Washington	8840	1923966	2031236	09/90
614957	Merchants Bank	8800	1694902	1694902	09/90
758066	NBC Bank, San Antonio NA	7240	1137517	2308082	06/90
1430912	Alliance Bank	380	1122637	1122637	06/89
996363	MBank Alamo NA	7240	856970	22339037	03/89
144258	Mbank, Austin NA	640	808506	22339037	03/89
531111	First New York Bank for Business	5600	665920	665920	12/92
938309	Capital Bank and Trust Co.	1120	507018	507018	12/90
7353	First City Texas, Austin NA	640	481653	13070185	12/92
61739	American Bank and Trust Co.	760	431764	490829	09/90
235802	Broadway Bank & Trust Co.	875	426315	426315	03/92
226453	Texas AMN Bank, Galleria NA	3360	398106	5817439	09/89
822707	University Bank NA	1120	398005	398005	06/91
145152	Louisiana Bank and Trust	7680	393895	393895	03/89
329251	First State Bank	40	385057	674591	03/89
678669	First City Texas, San Antonio NA	7240	369708	13070185	12/92
377056	Mbank Midcities NA	2800	369520	22339037	03/89
971904	Boston Trade Banks	1120	352142	352142	06/91
632102	Merchants Bank & Trust Co.	8040	319839	319839	03/91
744153	Bank of the Hills	640	303432	303432	03/91
637367	Mbank Jefferson County NA	840	289986	22339037	03/89

B. Holdout Sample (n = 25)

Bank ID	Name	MSA	Bank Assets[a]	BHC Assets[a]	Failure Date
913904	Bank of New England NA	1120	14292638	32635459	03/91
2509	Connecticut Bank & Trust Co. NA	3280	9735030	32635459	03/91
3850	First City Texas, Houston NA	3360	4829341	13070185	12/92
539201	CityTrust	1160	2240996	2240996	09/91
803014	First National Bank of Toms River	5190	1846892	1960839	06/91
546656	First City Texas, Dallas NA	1920	1734959	13070185	12/92
351038	First American Bank and Trust	8960	1566575	1566575	12/89
145303	Maine National Bank	6400	1255527	32635459	03/91
312253	Mbank, Fort Worth NA	2800	906754	22339037	03/89
798763	Independence Bank	4480	707501	707501	03/92
164658	First City Texas, Beaumont NA	840	707500	13070185	12/92
958156	First City Texas, Corpus Christi NA	1880	550563	13070185	12/92
944953	Metro North State Bank	3760	547220	547220	12/92
878508	Home National Bank of Milford	1120	520112	520112	06/90
518608	Nashua Trust Co.	5350	481543	610279	12/91
928654	First City Texas, El Paso NA	2320	475720	13070185	12/92
967121	Madison National Bank	8840	458192	784165	06/91
778701	Guaranty First Trust Co.	1120	440311	440311	12/92
656201	Guardian National Bank NA	5380	429529	429529	06/89
225410	Community Nat. Bank and Trust Co., New York	5600	396322	396322	12/91
234953	First City Texas, Bryan/College Station NA	1260	389845	13070185	12/92
694708	Coolidge Bank & Trust Co.	1120	345049	345049	12/91
483461	First City Texas, Tyler NA	8640	316760	13070185	12/92
670609	Landmark Bank	3280	277760	387723	03/91
886455	MBank Longview NA	4420	235254	22339037	03/89

[a]Asset sizes are based on the quarter immediately preceding the date of failure.

Table 2. Independent Variables and Basic Statistics[a]

		One Year Prior: Original Sample		One Year Prior: Holdout Sample		Two Years Prior: Original Sample		Two Years Prior: Holdout Sample	
x_j	Variable Description	Non-failed	Failed	Non-failed	Failed	Non-failed	Failed	Non-failed	Failed
1	Total assets (millions)	1,453	1,756	1,400	1,827	1,422	1,732	1,422	1,732
2	Bank holding co. (BHC) total assets (millions)	24,903	7,958*	15,973	9,800	1,829	8,690	18,295	8,690
3	Maximum change in X1/mean X1	0.03	0.05	0.03	0.04	0.66	0.55	0.66	0.55
4	Maximum change in X1/mean change X1	0.73	-0.11	1.49	6.43	1.29	0.81	1.29	0.81
5	Net interest income/total assets	0.03	0.02*	0.03	0.02*	0.03	0.03	0.03	0.03
6	Max change in X5/mean X5	0.19	0.13	0.12	-0.09	0.26	0.22	0.26	0.22
7	Net income after taxes/total assets	-0.001	-0.04*	0.01	-0.04*	0.001	-0.02*	0.001	-0.02*
8	Maximum change in X7/mean X7	-0.24	-15.75	-0.25	-1.08	-0.21	-4.14*	-0.021	-4.14*
9	Total equity/total assets	0.06	0.03*	0.06	0.04*	0.06	0.05*	0.06	0.05*
10	Maximum change in X9/mean X9	0.07	0.25	0.08	0.15*	0.12	0.32*	0.12	0.32*
11	Allowance for loan losses/total assets	0.01	0.02*	0.01	0.02	0.01	0.02	0.01	0.01
12	Max change in X11/mean X11	0.16	0.26	0.13	0.19	0.59	0.34	0.59	0.34
13	Provision for loan losses/total assets	0.01	0.01*	0.01	0.01*	0.01	0.01*	0.01	0.01*
14	Maximum change in X13/mean X13	1.05	1.39	1.13	1.15	1.09	0.92	1.10	0.92
15	Net loan charge-offs/total assets	0.01	0.03*	0.01	0.02*	0.01	0.01	0.01	0.01
16	Maximum change in X15/mean X15	1.06	0.87	1.07	0.67	1.39	0.78*	1.39	0.78*
17	Maximum change in loans past due at least 90 days/mean of numerator	0.79	0.63	0.64	0.73	1.03	0.68	1.03	0.68
18	Total securities/total assets	0.17	0.15	0.19	0.14	0.15	0.14	0.15	0.14
19	Maximum change in X18/mean X18	0.09	0.17	0.09	0.15	0.76	0.63	0.76	0.63
20	Nondeposit liabilities/total liabilities	-0.02	-0.17*	-0.13	-0.17	0.02	-0.06	0.02	-0.07
21	Maximum change in X20/mean X20	1.69	-0.65	0.37	2.66	0.03	-1.57*	0.03	-1.57*
22	Certificates of deposit/total deposits	0.32	0.43*	0.37	0.40	0.37	0.43*	0.37	0.43*
23	Maximum change in X22/mean X22	0.07	0.07	0.08	0.03	0.24	0.10	0.24	0.10
24	Total loans and leases/total assets	0.56	0.63	0.61	0.63	0.59	0.64	0.59	0.64
25	Maximum change in X24/mean X24	0.06	0.08	0.07	0.10	0.12	0.05*	0.12	0.05*
26	Sum of key asset accounts/total assets (quantity squared as in a HHI index)	0.04	0.05	0.05	0.05	0.04	0.05	0.04	0.06
27	Maximum change in X26/mean X26	0.18	0.26	0.16	0.23	0.42	0.53	0.42	0.54
28	Bank total assets/BHC total assets	0.50	0.60	0.53	0.54	0.41	0.59	0.45	0.60

[a]Variables are calculated at their levels in the quarter prior to failure for the failed (n = 50) and matched-pair non-failed (n = 50) banks. Numerous variables are calculated in change form. For these change variables the maximum difference between quarters over the previous four quarters is divided either by the mean level of the variable or the mean change between quarters as indicated in the denominator. Asterisks indicate a significant difference at the 10 percent level between non-failed and failed banks' mean values for a particular variable.

Table 3. Results of Stepwise Logit Model[a]

A. One Year Prior to Failure

1. Sample with n = 50 (25 failed banks and 25 non-failed banks)
   Model: -3.38 - 1,854.80 X9 + 536.90 X11 + 9.46 X16 + 171.40 X22 + 22.89 X28

Original Sample

Actual:	Predicted: Fail	Predicted: Non-fail
Fail	25	0
Non-fail	0	25

Percentage Correct: 100.0%

Holdout Sample

Actual:	Predicted: Fail	Predicted: Non-fail
Fail	14	11
Non-fail	4	21

Percentage Correct: 70.0%

2. Sample with n = 30 (15 failed banks and 15 non-failed banks)
   Model: -168.40 - 1,527.50 X9 + 556.60 X22 + 153.70 X27

Original Sample

Actual:	Predicted: Fail	Predicted: Non-fail
Fail	15	0
Non-fail	0	15

Percentage Correct: 100.0%

Holdout Sample

Actual:	Predicted: Fail	Predicted: Non-fail
Fail	8	7
Non-fail	1	14

Percentage Correct: 73.3%

3. Sample with n = 20 (10 failed banks and 10 non-failed banks)
   Model: -22.12 + 52.38 X22 + 15.86 X27

Original Sample

Actual:	Predicted: Fail	Predicted: Non-fail
Fail	9	1
Non-fail	1	9

Percentage Correct: 90.0%

Holdout Sample

Actual:	Predicted: Fail	Predicted: Non-fail
Fail	5	5
Non-fail	3	7

Percentage Correct: 60.0%

Table 3, continued

B. Two Years Prior to Failure

1. Sample with n = 50 (25 failed banks and 25 non-failed banks)
Model: 29.69 – 634.20 X7 – 1,063.00 X9 – 24.98 X20 – 5.44 X21 + 141.10 X26 – 37.38X28

Original Sample

Actual:	Predicted: Fail	Predicted: Non-fail
Fail	25	0
Non-fail	2	23

Percentage Correct: 96.0%

Holdout Sample

Actual:	Predicted: Fail	Predicted: Non-fail
Fail	11	14
Non-fail	6	19

Percentage Correct: 60.0%

2. Sample with n = 30 (15 failed banks and 15 non-failed banks)
Model: -34.71 - 1,434.30 X7 - 2.09 X8 + 2,601.20 X11 - 3,414.50 X15 – 11.88 X16 + 51.55 X28

Original Sample

Actual:	Predicted: Fail	Predicted: Non-fail
Fail	15	0
Non-fail	0	15

Percentage Correct: 100.0%

Holdout Sample

Actual:	Predicted: Fail	Predicted: Non-fail
Fail	7	8
Non-fail	2	13

Percentage Correct: 66.7%

3. Sample with n = 20 (10 failed banks and 10 non-failed banks)
Model: -38.70 - 5.25E-7 X2 - 1,541.20 X7 + 44.00 X10 + 727.10 X26

Original Sample

Actual:	Predicted: Fail	Predicted: Non-fail
Fail	10	0
Non-fail	0	10

Percentage Correct: 100.0%

Holdout Sample

Actual:	Predicted: Fail	Predicted: Non-fail
Fail	6	4
Non-fail	1	9

Percentage Correct: 75.0%

---

[a]The 28 variables listed in table 2 were entered as eligible for the stepwise logistic procedure. Final maximum likelihood estimates of the parameters were not possible for most models because of sample size limitations. Results are shown based on the last maximum likelihood iteration.

Table 4. Cutpoints for Variables in the Trait Recognition Models (n = 50)

	One Year Prior		Two Years Prior			
	Stage 1		Stage 1		Stage 2	
Variable	Left	Right	Left	Right	Left	Right
X1	-0.45	0.31	-0.36	0.25	-0.17	0.01
X2	-0.51	0.22	-0.52	0.53	-0.45	0.92
X3	-0.24	0.29	-.053	0.01	0.00	0.01
X4	-0.31	0.14	-0.63	0.01	-0.68	0.01
X5	0.00	0.01	-0.37	0.07	-0.01	0.89
X6	-0.21	0.01	-0.20	0.01	-0.10	0.01
X7	0.00	0.08	-0.28	0.18	-0.47	0.43
X8	0.00	0.19	-0.08	0.32	-0.06	0.02
X9	-0.03	0.05	-0.03	0.01	-0.09	0.01
X10	-0.13	0.01	-0.02	0.06	0.00	0.49
X11	-0.17	0.01	-0.53	0.01	-0.31	0.01
X12	-0.49	0.19	-0.54	0.69	-0.52	0.46
X13	-0.58	0.01	-0.51	0.27	-0.61	0.20
X14	-0.01	0.99	-0.67	0.67	-0.07	0.03
X15	-0.45	0.44	-0.27	0.10	-0.49	0.01
X16	-0.46	0.01	0.00	0.10	-0.56	0.13
X17	-0.50	1.89	0.00	0.04	-0.73	0.07
X18	0.00	0.01	-0.90	0.19	0.00	0.01
X19	-0.48	0.52	-0.34	1.19	-0.35	0.68
X20	-0.31	0.01	-0.49	0.27	-1.14	0.90
X21	-0.10	0.28	0.00	0.10	-0.14	0.01
X22	0.00	0.01	-0.17	0.31	0.00	0.53
X23	-0.14	0.01	-0.07	0.39	0.00	0.23
X24	-0.01	0.21	-0.59	0.59	-0.56	0.58
X25	-0.78	0.48	0.00	0.30	0.00	0.01
X26	-0.71	0.37	0.00	0.02	0.00	0.01
X27	-0.49	0.17	-0.17	0.61	-0.26	0.77
X28	-0.52	0.69	-0.54	0.37	0.00	0.33

Table 5. Trait Recognition Model Features (n = 50): One Year and Two Years Prior to Failure

A. One Year Prior to Failure (25 failed banks and 25 non-failed banks)

Region: L | M | U; Coding: 00 | 01 | 11

1. Safe Features (35)

p	q	r	P	Q	R	X_j	(j=1, …, 28)		Region		
2	15	18	0	1	1	X1	X8	X9	L	U	MU
2	15	20	0	1	0	X1	X8	X10	L	U	L
2	17	22	0	1	0	X1	X9	X11	L	U	L
9	18	30	1	1	0	X5	X9	X15	U	MU	L
13	15	18	1	1	1	X7	X8	X9	U	U	U
13	17	21	1	1	0	X7	X9	X11	U	U	LM
15	17	18	1	1	1	X8	X9	X9	U	U	MU
15	17	21	1	1	0	X8	X9	X11	U	U	LM
15	17	22	1	1	0	X8	X9	X11	U	U	L
15	17	49	1	1	1	X8	X9	X25	U	U	U
15	18	19	1	1	0	X8	X9	X10	U	MU	LM
15	18	24	1	1	0	X8	X9	X12	U	MU	L
15	18	28	1	1	0	X8	X9	X14	U	MU	L
15	18	30	1	1	0	X8	X9	X15	U	MU	L
15	18	42	1	1	0	X8	X9	X21	U	MU	L
15	18	51	1	1	1	X8	X9	X26	U	MU	U
15	19	19	1	0	0	X8	X10	X10	U	LM	LM
15	19	28	1	0	0	X8	X10	X14	U	LM	L
15	19	30	1	0	0	X8	X10	X15	U	LM	L
15	20	24	1	0	0	X8	X10	X12	U	L	L
15	20	28	1	0	0	X8	X10	X14	U	L	L
15	20	30	1	0	0	X8	X10	X15	U	L	L
15	20	49	1	0	1	X8	X10	X25	U	L	U
15	20	51	1	0	1	X8	X10	X26	U	L	U
15	24	30	1	0	0	X8	X12	X15	U	L	L
15	28	43	1	0	0	X8	X14	X22	U	L	LM
17	21	28	1	0	0	X9	X11	X14	U	LM	L
17	21	30	1	0	0	X9	X11	X15	U	LM	L
17	22	28	1	0	0	X9	X11	X14	U	L	L
17	22	30	1	0	0	X9	X11	X15	U	L	L
17	22	49	1	0	1	X9	X11	X25	U	L	U
24	30	51	0	0	1	X12	X15	X26	L	L	U
24	30	54	0	0	0	X12	X15	X27	L	L	L
28	30	38	0	0	0	X14	X15	X19	L	L	L
28	30	51	0	0	1	X14	X15	X26	L	L	U

Region: L | M | U; Coding: 00 | 01 | 11

2. Unsafe Features (36)

p	q	r	P	Q	R	X_j	(j=1, …, 28)		Region		
1	7	11	1	1	1	X1	X4	X6	U	U	U
1	11	11	1	1	1	X1	X6	X6	U	U	U
1	11	34	1	1	0	X1	X6	X17	U	U	L
1	11	42	1	1	0	X1	X6	X21	U	U	L
1	16	16	1	0	0	X1	X8	X8	U	L	L
1	25	25	1	1	1	X1	X13	X13	U	U	U
2	16	16	0	0	0	X1	X8	X8	L	L	L
7	11	11	1	1	1	X4	X6	X6	U	U	U
7	11	16	1	1	0	X4	X6	X8	U	U	L
7	11	25	1	1	1	X4	X6	X13	U	U	U
7	11	34	1	1	0	X4	X6	X17	U	U	L
7	11	42	1	1	0	X4	X6	X21	U	U	L
7	16	16	1	0	0	X4	X8	X8	U	L	L
7	16	25	1	0	1	X4	X8	X13	U	L	U
7	16	34	1	0	0	X4	X8	X17	U	L	L
7	18	18	1	0	0	X4	X9	X9	U	L	L
7	25	25	1	1	1	X4	X13	X13	U	U	U
7	25	34	1	1	0	X4	X13	X17	U	U	L
7	29	29	1	1	1	X4	X15	X15	U	U	U
9	9	9	0	0	0	X5	X5	X5	LM	LM	LM
11	16	16	1	0	0	X6	X8	X8	U	L	L
11	25	25	1	1	1	X6	X13	X13	U	U	U
11	29	29	1	1	1	X6	X15	X15	U	U	U
11	34	42	1	0	0	X6	X17	X21	U	L	L
16	16	16	1	0	0	X8	X8	X8	MU	L	L
16	34	34	0	0	0	X8	X17	X17	L	L	L
18	18	18	0	0	0	X9	X9	X9	L	L	L
19	19	19	1	1	1	X10	X10	X10	U	U	U
21	21	21	1	1	1	X11	X11	X11	U	U	U
25	25	25	1	1	1	X13	X13	X13	U	U	U
25	29	29	1	1	1	X13	X15	X15	U	U	U
25	34	34	1	0	0	X13	X17	X17	U	L	L
29	29	29	1	1	1	X15	X15	X15	U	U	U
29	34	34	1	0	0	X15	X17	X17	U	L	L
29	42	42	1	0	0	X15	X21	X21	U	L	L
42	43	43	0	1	1	X21	X22	X22	L	U	U

Table 5, continued

B. Two Years Prior to Failure (25 failed banks and 25 non-failed banks)

1. Stage 1-- Safe Features (9)

L	M	U	
00	01	11	-Coding

(L, M, U -Region)

Program Coding						Variables and Location					
p	q	r	P	Q	R	X_J	(j=1,…,28)		Region		
4	13	22	0	1	0	X2	X7	X11	L	U	L
13	14	22	1	1	0	X7	X7	X11	U	MU	L
13	15	30	1	1	9	X7	X8	X15	U	LM	L
13	44	44	1	0	1	X7	X22	X22	U		
13	51	51	1	0	0	X7	X26	X26	U	L	MU
17	17	17	1	1	1	X9	X9	X9	U	U	U
20	26	26	0	0	0	X10	X13	X13	L	L	L
22	26	26	0	0	0	x11	X13	X13	L	L	L
22	51	51	0	0	0	x11	X26	X26	L	LM	LM

2. Stage 1--Unsafe Features(5)

L	M	U	
00	01	11	-Coding

(L, M, U -Region)

Program Coding						Variables and Location					
p	Q	r	P	Q	R	X_j	(j=1,…,28)		Region		
17	17	17	0	0	0	X9	X9	X9	LM	LM	LM
17	18	35	0	0	1	X9	X9	X18	LM	L	U
17	18	38	0	0	0	X9	X9	X19	LM	L	L
21	35	35	1	1	1	X11	X19	X19	U	L	L
21	38	38	1	0	0	X11	X19	X19	U	L	L

3. Stage 2--Safe Features (18)

Program Coding						Variables and Location					
p	q	r	P	Q	R	X_j	(j=1,…,28)		Region		
4	19	21	0	0	0	X2	X10	X11	L	LM	LM
7	13	38	1	1	0	X4	X7	X19	U	U	L
7	13	39	1	1	1	X4	X7	X20	U	U	U
7	19	41	1	0	1	X4	X10	X21	U	LM	U
13	14	39	1	1	1	X7	X7	X20	U	MU	U
13	18	51	1	1	9	X7	X9	X26	U	MU	LM
13	19	39	1	0	1	X7	X10	X26	U	LM	LM
13	19	51	1	0	0	X7	X10	X26	U	LM	LM
13	39	51	1	0	0	X7	X20	X26	U	U	MU
13	51	51	1	1	0	X7	X26	X26	U	U	MU
15	20	39	1	0	1	X8	X10	X20	U	L	U
15	20	41	1	0	1	X8	X10	X21	U	L	U
15	33	39	1	1	1	X8	X17	X20	U	U	U
17	38	44	1	0	0	X9	X19	X22	U	L	L
17	39	44	1	1	0	X9	X20	X22	U	U	L
17	41	44	1	0	0	X9	X21	X22	U	U	L
17	44	44	1	0	0	X9	X22	X22	U	L	L
17	44	54	1	0	0	X9	X22	X27	U	L	L

4. Stage 2--Unsafe Features (14)

Program Coding						Variables and Location					
p	q	r	P	Q	R	X_j	(j=1,…,28)		Region		
4	29	49	0	1	0	X2	X15	X25	L	U	LM
5	6	21	0	0	1	X3	X3	X11	LM	L	U
5	10	47	0	0	1	X3	X5	X24	LM	L	U
10	17	17	0	0	0	X5	X9	X9	L	LM	LM
10	18	18	0	0	1	X5	X9	X9	L	L	LM
10	18	29	0	0	1	X5	X9	X15	L	L	U
10	18	32	0	0	0	X5	X9	X16	L	L	L
10	18	46	0	0	0	X5	X9	X23	L	L	L
10	21	21	0	1	1	X5	X11	X11	L	U	U
10	21	32	0	1	0	X5	X11	X16	L	U	L
10	29	32	0	1	0	X5	X15	X16	L	U	L
18	29	32	0	1	0	X9	X15	X16	L	U	L
29	46	49	1	0	0	X15	X23	X25	U	LM	LM
29	47	47	1	1	1	X15	X24	X24	U	U	U

Table 6. Trait Recognition Model Results

A. One Year Prior to Failure

1. Sample with n = 50 (25 failed banks and 25 non-failed banks)
   Model: 35 safe features and 36 unsafe features (see table 5)
   Safe features – 80% minimum non-failed banks and 30% maximum failed banks
   Unsafe features – 75% minimum failed banks and 40% maximum non-failed banks

Original Sample (Predicted:)

Actual:	Fail	Non-fail
Fail	25	0
Non-fail	0	25

Percentage Correct: 100.0%

Holdout Sample (Predicted:)

Actual:	Fail	Non-fail
Fail	23	2
Non-fail	8	17

Percentage Correct: 80.0%

2. Sample with n = 30 (15 failed banks and 15 non-failed banks)
   Model: 9 safe features and 13 unsafe features
   Safe features – 80% minimum non-failed banks and 20% maximum failed banks
   Unsafe features – 72% minimum failed banks and 25% maximum non-failed banks

Original Sample (Predicted:)

Actual:	Fail	Non-fail
Fail	15	0
Non-fail	0	15

Percentage Correct: 100.0%

Holdout Sample (Predicted:)

Actual:	Fail	Non-fail
Fail	11	4
Non-fail	2	13

Percentage Correct: 80.0%

3. Sample with n = 20 (10 failed and 10 non-failed banks)
   Model: 6 safe features and 6 unsafe features
   Safe features – 70% minimum non-failed banks and 30% maximum failed banks
   Unsafe features – 80% minimum failed banks and 20% maximum non-failed banks

Original Sample (Predicted:)

Actual:	Fail	Non-fail
Fail	10	0
Non-fail	0	10

Percentage Correct: 100.0%

Holdout Sample (Predicted:)

Actual:	Fail	Non-fail
Fail	10	0
Non-fail	1	9

Percentage Correct: 95.0%

Table 6, continued

B. Two Years Prior to Failure

1. Sample with n = 50 (25 failed banks and 25 non-failed banks)
   Models: Stage 1 – 9 safe features and 5 unsafe features (see table 5)
   Stage 2 – 18 safe features and 14 unsafe features (see table 5)
   Safe features – 75% (Stage 1) or 79% (Stage 2) minimum non-failed banks and 25% (Stage 1 and Stage 2) maximum failed banks
   Unsafe features – 75% (Stage 1) or 79% (Stage 2) maximum failed banks and 25% (Stage 1 and Stage 2) minimum non-failed banks

Original Sample

Predicted: (Stage 1/Stage 2)

Actual:	Fail	Non-fail
Fail	10/15=25	0/0=0
Non-fail	0/0=0	12/13=25

Percentage Correct: 100.0%

Holdout Sample

Predicted: (Stage 1/Stage 2)

Actual:	Fail	Non-fail
Fail	8/9=17	4/4=8
Non-fail	5/2=7	7/11=18

Percentage Correct: 70.0%

2. Sample with n = 30 (15 failed banks and 15 non-failed banks)
   Models: 11 safe features and 10 unsafe features
   Safe features – 80% minimum non-failed banks and 40% maximum failed banks
   Unsafe features – 85% minimum failed banks and 25% maximum non-failed banks

Original Sample

Predicted:

Actual:	Fail	Non-fail
Fail	15	0
Non-fail	0	15

Percentage Correct: 100.0%

Holdout Sample

Predicted:

Actual:	Fail	Non-fail
Fail	10	6
Non-fail	5	9

Percentage Correct: 63.3%

3. Sample with n = 20 (10 failed banks and 10 non-failed banks)
   Model: 9 safe features and 12 unsafe features
   Safe features: 70% minimum non-failed banks and 40% maximum failed banks
   Unsafe features: 70% minimum failed banks and 10% maximum non-failed banks

Original Sample

Predicted:

Actual:	Fail	Non-fail
Fail	10	0
Non-fail	0	10

Percentage Correct: 100.0%

Holdout Sample

Predicted:

Actual:	Fail	Non-fail
Fail	8	2
Non-fail	4	6

Percentage Correct: 70.0%

Table 7. Weighted Efficiency Scores: Logit and Trait Recognition Models[a]

	FCC	PF	AF	CC	WE
A. Full holdout sample (n=50)					
One year prior:					
Logit	14	18	25	70.0	30.5
Trait recognition	23	31	25	80.0	54.6
Two years prior:					
Logit	11	17	25	60.0	17.1
Trait recognition	17	24	25	70.0	33.7
B. Holdout sample (n=30)					
One year prior:					
Logit	8	9	15	73.3	34.7
Trait recognition	11	13	15	80.0	49.6
Two years prior:					
Logit	7	9	15	66.7	24.2
Trait recognition	10	15	15	63.3	28.1
C. Holdout sample (n=20)					
One year prior:					
Logit	5	8	10	60.0	24.0
Trait recognition	10	11	10	95.0	86.4
Two years prior:					
Logit	6	7	10	75.0	31.5
Trait recognition	8	12	10	70.0	37.3

[a]The weighted efficiency score is calculated as follows: WE = (FCC/PF)*(FCC/AF)*CC, where FCC = the number of failed banks classified correctly, PF = the number of banks predicted to fail, AF = the number of banks that actually failed, and CC = the percentage of banks correctly classified.

Table 8. Logit Predicted Probabilities for Large Failed Bank Holdout Samples[a]
(ranked from largest to smallest in asset size one year prior to failure[b])

	Holdout Sample (n = 25)		Holdout Sample (n = 15)		Holdout Sample (n = 10)	
Bank ID	One Year Prior	Two Years Prior	One Year Prior	Two Years Prior	One Year Prior	Two Years Prior
913904	1.000	0.000*	1.000	0.000*	0.729	0.000*
2509	1.000	0.000*	1.000	0.000*	0.573	0.000*
3850	1.000	1.000	1.000	1.000	0.389*	1.000
539201	1.000	1.000	0.000*	1.000	0.034*	1.000
546656	1.000	0.000*	0.000*	1.000	0.012*	1.000
803014	0.000*	0.001*	0.000*	0.000*	0.820	0.000*
351038	1.000	0.002*	1.000	1.000	0.923	1.000
145303	0.000*	0.000*	0.000*	0.000*	0.036*	0.000*
312253	1.000	1.000	1.000	0.000*	1.000	1.000
798763	0.000*	1.000	0.000*	1.000	0.068*	1.000
958156	0.000*	0.000*	0.000*	0.000*		
164658	0.000*	0.000*	1.000	0.000*		
944953	1.000	1.000	1.000	1.000		
778701	1.000	1.000	1.000	0.973		
878508	1.000	0.000*	0.000*	0.000*		
928654	0.000*	1.000				
967121	0.000*	0.000*				
518608	0.000*	1.000				
694708	0.736	0.998				
225410	0.001*	1.000				
234953	0.000*	0.000*				
483461	0.000*	0.000*				
656201	1.000	1.000				
670609	1.000	0.063*				
886455	1.000	0.484*				

[a]An asterisk denotes a prediction error for a bank in a particular model and holdout sample.

[b]For the asset size of each bank in the quarter immediately preceding the failure date, see table 1.

Table 9. Trait Recognition Voting Results for Holdout Samples of Large Failed Banks [a]
(ranked from largest to smallest by asset size one year prior to failure[b])

	Holdout Sample (n = 25)				Holdout Sample (n = 15)				Holdout Sample (n = 10)			
	One Year Prior		Two Years Prior		One Year Prior		Two Years Prior		One Year Prior		Two Years Prior	
Bank ID	Safe Votes	Unsafe Votes	Safe Votes	Unsafe Votes	Safe Votes	Unsafe Votes	Safe Votes	Unsafe Votes	Safe Votes	Unsafe Votes	Safe Votes	Unsafe Votes
913904	12	26	5 (5)	5 (13)	0	9	3	3	9	5 [c]	9	7 [c]
2509	7	25	6	2*	0	6	8	0*	0	0	1	1
3850	5	34	0 (0)	4 (7)	0	6	0	7	0	4	0	11
539201	0	35	0 (3)	5 (14)	0	6	0	10	0	5	0	6
546656	1	34	0	3	0	5	4	6	0	3	0	2
803014	32	21*	9 (18)	0 (1)*	9	2*	10	0*	5	5	8	0*
351038	10	12	5 (2)	5 (7)	1	2	2	3	1	1	4	7
145303	15	24	8	2*	1	6	7	0*	5	5	8	0*
312253	4	36	0 (1)	1 (10)	0	9	1	9	0	6	0	9
798763	5	22	0 (3)	5 (9)	7	4*	2	3	4	4	5	5
958156	9	25	4	2 [c]	6	1*	6	0*				
164658	12	13	8	0*	5	4*	7	0*				
944953	0	35	3 (4)	2 (4)	0	8	5	6				
778701	0	36	1 (0)	2 (6)	0	8	2	7				
878508	16	24	4	2 [c]	2	9	9	6*				
928654	1	12	3	5								
967121	24	9*	7	1*								
518608	4	31	0 (5)	5(4)*								
694708	4	20	4	2 [c]								
225410	14	23	6 (1)	3 (3)								
234953	1	13	6 (6)	0 (0)*								
483461	5	12	4	2 [c]								
656201	1	17	5 (8)	3 (2)*								
670609	0	29	5	2 [c]								
886455	0	22	0	3								

[a] An asterisk denotes a prediction error for a bank in a particular model and holdout sample based on the number of safe and unsafe votes. In the two-years-prior results for the n = 50 sample, a two-stage model was used, and stage 1 results (and stage 2 results in parentheses) are shown.

[b] For the asset size of each bank in the quarter immediately preceding the failure date, see table 1.

[c] These banks had more safe votes than unsafe votes but were still predicted to fail. The reason that they were not misclassified as safe banks is that these cells in the voting matrix (e.g., 4 safe and 2 unsafe votes) were identified as unsafe in the procedure used to catalog safe, unsafe, mixed, and unclassified cells.