A Pattern Recognition Approach to Early Warning Systems in Commercial Banking

James Kolari
Professor of Finance
Texas A&M University
Finance Department
College Station, TX 77843
U.S.A.
Telephone: 409-845-4803

Michele Caputo
Professor of Geophysics
Instituto di Fisica della Università di Roma
Italy

Drew Wagner
Doctoral Candidate
Texas A&M University
Finance Department

Summary

The present study comparatively examines the classification and predictive ability of pattern recognition relative to common statistical approaches, such as multiple discriminant analysis (MDA) and logit analysis. The latter empirical methods have been applied by bank regulators as early warning systems (EWSs) to monitor bank condition based on financial ratios between on-site examinations. The pattern recognition technique applied in the present study differs from previous EWS models in that all possible interactions of the independent variables taken one, two, and three at a time are evaluated for their potential usefulness in discrimination. The results of the comparative analyses of failed and nonfailed U.S. commercial banks using 1984 and 1985 Call Report data indicate that pattern recognition generally outperformed MDA and logit EWS models, in many cases by a considerable margin. In light of the record numbers of bank failures in recent years, pattern recognition methods appear to offer another regulatory tool that could be valuable in identifying failing banks prior to collapse. Moreover, it can be applied in other disciplines, in which discrimination between different groups of observations is desired (e.g., automated credit scoring methods in consumer and business applications, insurance risk evaluation of individuals and firms according to different classes, etc.).

La reconnaissance des modèles appliquée aux systèmes d'alerte avancée pour les banques commerciales

James Kolari
Professeur de finances
Texas A&M University
Finance Department
College Station, TX 77843
U.S.A.
Téléphone : 409-845-4803

Michele Caputo
Professeur de géophysique
Institut de physique de l'université de Rome
Italie

Drew Wagner
Candidat au doctorat
Texas A&M University
Finance Department

Résumé

La présente étude examine de manière comparative la classification et la valeur de prédiction de la reconnaissance des modèles de diverses méthodes statistiques courantes, telles que l'analyse multivariables (AMV) et l'analyse logit. Ces méthodes empiriques ont été appliquées par les régulateurs bancaires en tant que système d'alerte avancée (SAA) pour surveiller la situation des banques sur la base des ratios financiers, entre les audits sur place. La technique de reconnaissance des formes appliquée dans la présente étude diffère des modèles de SAA antérieurs en ce que toutes les interactions possibles des variables indépendantes prises une, deux et trois à la fois sont évaluées pour déterminer leur utilité potentielle pour la discrimination. Les résultats des analyses comparatives des banques commerciales américaines ayant et n'ayant pas fait faillite, sur la base des données du Call Report pour 1984 et 1985, indiquent que la reconnaissance des formes est généralement supérieure, et souvent dans une mesure considérable, aux modèles de SAA par AMV et par analyse logit. Compte tenu du nombre record de faillites bancaires survenues ces dernières années, les méthodes de reconnaissance des formes semblent constituer un instrument réglementaire supplémentaire qui pourrait être utile pour identifier les banques en difficultés avant leur faillite effective. En outre, elles peuvent être appliquées dans d'autres disciplines, dans lesquelles il est souhaitable d'effectuer une discrimination entre différents groupes d'observations (tels que les méthodes automatisées d'évaluation du crédit appliquées aux consommateurs et aux entreprises, l'évaluation des risques d'assurance de personnes et d'entreprises selon les différentes catégories, etc.).

A Pattern Recognition Approach to Early Warning Systems in Commercial Banking

Throughout the 1980s and early 1990s there were record numbers of post-Depression failed U.S. depository institutions, including commercial banks, mutual savings banks, credit unions, and (particularly) savings and loan associations (S&Ls). Many failed S&Ls were subsequently absorbed into the commercial banking industry. However, commercial banks have not been immune to problems, as a total of 122 U.S. commercial banks failed in 1992 with assets exceeding $40 billion. As the number of bank failures has escalated, so has the cost to the government of dealing with them. In an effort to minimize the costs associated with bank failure, bank regulators and managers seek to act quickly in order to either prevent bank failure or lower the cost of failure. One tool utilized by federal regulatory agencies in the U.S. and other countries is *early warning systems* (EWSs) that attempt to predict potential problems with commercial banks and other depository institutions [see Thomson (1991)]. Institutions that trigger the EWS system become subject to increased regulatory attention. In this way EWSs provide lead time to improve the allocation of scarce examiner resources, allow timely supervisory actions, and ultimately reduce failure costs.

Recent work by Espahbodi(1991) has shown that logit models tend to outperform multiple discriminant (MDA) models as EWSs. While a large body of empirical evidence using these statistical models[1] has proven their effectiveness as EWSs, Frydman *et al.* (1985) have noted that nonparametric classification procedures could be an alternative (and little studied) approach worthy of testing. They employed a recursive partitioning

technique, which is based on a regression tree,[2] to predict failed nonfinancial firms. The results confirmed their hypothesis that nonparametric techniques have merit as EWSs, as the recursive partitioning model outperformed a MDA model.[3]

The present study extends work on nonparametric EWSs by applying a pattern recognition (PR) technique to samples of failed and nonfailed U.S. commercial banks. Unlike previous EWSs, the technique recodes financial ratio data into binary strings and then examines all possible combinations of the variables taken one, two, and three at a time for their usefulness in discriminating failed and nonfailed banks. Because these combinations are known as traits, we will refer to the method as *trait recognition* (TR) to distinguish it from other PR techniques. Previously, this TR method has been applied to a variety of identification problems in the sciences, including earthquake prediction (Gelfand *et al.*, 1972; and Benavidez and Caputo, 1988), uranium detection (Briggs and Press, 1977), and oil exploration (Bongard *et al.*, 1966). However, to our knowledge, it has not been employed in the field of business and economics. In this study, the classification and prediction power of a TR model, in addition to a hybrid model using both TR and MDA (recommended by Frydman *et al.*), are compared to both MDA and logit models. The results of these comparative analyses show that TR outperformed both MDA and logit models in most instances. Indeed, the predictive ability of TR on holdout samples exceeds most published studies on the general subject of predicting firm failure. We conclude that TR is a potentially useful EWS that could aid regulators, bankers, investors, and others in evaluating bank condition.

Section I provides background discussion of pattern recognition techniques and

describes the construction of our trait recognition (TR) EWS model. Section II gives details of our samples and financial variables, as well as an overview of MDA and logit models used in the comparative analyses. Section III reports the empirical results of comparing the classification and prediction power of TR, hybrid TR/MDA, MDA, and logit models. Section IV gives the conclusions and implications.

I. Trait Recognition Model Development

Background Discussion

Pattern recognition (PR) is a general term for computer-intensive processes that utilize input data to develop features (or attributes) which can be employed to discriminate between members of different groups. A broad interdisciplinary subject, PR spans many academic fields of study, including medicine, engineering, computer science (e.g., speech recognition), and the social sciences.[4] In most PR problems the following steps to system design are common:

> (1) quantifiably measuring the characteristics or traits of the observations and encoding this information;
> (2) preprocessing and extraction of distinctive features that represent common patterns of different groups of observations;
> (3) training or learning procedures on sample observations wherein arbitrary decision rules are initially applied and an iterative process is used to reach an optimum or satisfactory set of decision rules;
> (4) discrimination of observations in a holdout sample into different groups by the PR model.

Within this common framework, PR methods can be categorized as (1) heuristic, (2) mathematical (including statistical), and (3) linguistic (i.e., speech recognition).

In the present study we take a heuristic approach, which relies upon researcher

intuition and experience. While *ad hoc* in nature, heuristic PR methods are well suited to specialized tasks, such as our problem of identifying failing banks. In this regard, the large number of studies on bank failure prediction provides considerable experience in selecting variables.

The identification success of PR methods ultimately rests on the usefulness of the patterns found in the measurements. For example, in Briggs, Press, and Guberman (1977), a heuristic PR technique developed in Bongard *et al.* (1966) proved useful in identifying earthquake-prone areas. Remarking on the advantage of their heuristic PR model, Briggs *et al.* noted that, while certain variables give telltale signs of danger (e.g., the location of established faults):

> "...less obvious are the *interrelations* (emphasis added) of low-level seismic activity, topography, geothermal activity, recent volcanism, and so forth. We consider data on 45 such properties ... and search all combinations of one, two, and three of these properties to find characteristic traits ... Data are converted to yes/no or 1/0 binary code ... With a large number of properties or parameters, the impracticality of pattern recognition without machine analysis is clear." (1977, pp. 161-162)

Thus, the major strength of the Bongard *et al.* TR technique is the exhaustive search for interactions among a large number of independent variables that are useful in terms of identifying a particular group of observations.

In the present study we also employ the Bongard *et al.* algorithm; however, because the algorithm was custom programmed for geological studies, we were forced to create an entirely new program. In this regard, we developed a general version of the Bongard *et al.* algorithm that can be applied to virtually any identification problem.[5] Since the crucial aspects of the program are the quantification of all possible traits and then extraction of

traits that are common features of either failed or nonfailed banks, we refer to the algorithm as trait recognition (TR).

Description of Trait Recognition Procedure

For purposes of illustrating basic aspects of the trait recognition (TR) procedure, assume a researcher selects five nonfailed banks labeled with lower case letters a to e, and five failed banks labeled with upper case letters A to E. Based on experience, three financial ratios representative of bank financial condition are calculated for each bank one year prior to the collapse of the failed banks: net income/total assets (X_1), loan losses/total assets (X_2), and equity capital/total assets (X_3). These data are plotted as shown in Figure 1.

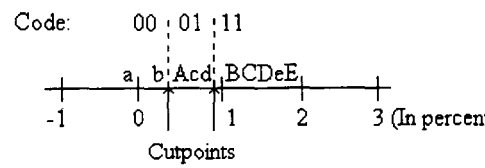
The next step is to choose two cutpoints for each financial ratio that segment the observations into three classes: (1) predominantly failed banks (coded 00), (2) a mixture of nonfailed and failed banks (coded 01), and (3) predominantly nonfailed banks (coded 11). The middle of the distribution (coded 01) is a grey area in which it is not clear that the ratio provides information that would discriminate between nonfailed and failed banks. As an example, Figure 1 shows a choice of cutpoints (marked with dashed lines) for our samples of five failed and five nonfailed banks using X_1, X_2, and X_3.[6] For X_1 all banks in the 00 segment are failed banks, the 01 segment is mixed with two failed banks and one nonfailed bank, and the 11 segment is dominated by four nonfailed banks and one failed bank. It would be possible to move the cutpoint for the 11 segment to a position just to the right of the failed bank e, in which case only nonfailed bank E would be in this segment.

Figure 1

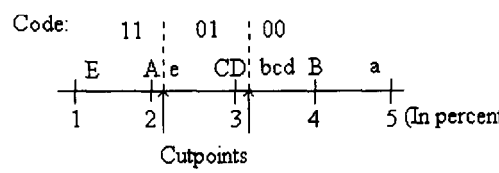
Plots of Financial Ratios for Samples
of Failed and Nonfailed Banks One
Year Prior to Failure

Samples — Failed banks: a, b, c, d, e
Nonfailed banks: A, B, C, D, E

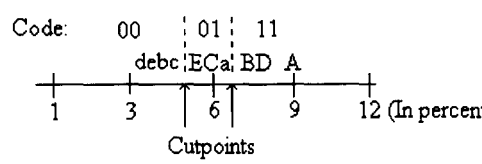
$X_1 = \frac{\text{net income}}{\text{total assets}}$



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



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



However, this restrictive selection of cutpoints would appear to ignore the fact that *most* nonfailed banks have net income/total assets ratios to the right of the position of nonfailed bank B. In our experience with TR, this restrictive approach to the choice of cutpoints was inferior to setting them in locations that captured the notion of dominance of one group or another. A similar logic holds for the choice of the cutpoints for X_2. For X_3 the cutpoints are set such that only failed banks are in segment 00, mixed banks in 01, and only nonfailed banks in 11; however, we should mention that this type of outcome was rare with larger bank samples.

Given the cutpoints for each variable, the data for each bank can be recoded. Based on Figure 1, the ten sample banks are coded into binary strings $A_1A_2...A_L$, where L is the length of the string and two digits describe each variable in the sequence $X_1X_2X_3$ as follows:

Failed banks		Nonfailed banks	
a	000001	A	011111
b	000000	B	110011
c	010000	C	110101
d	010000	D	110111
e	110100	E	111101

Notice that all binary strings have a different *pattern*, with the exception of failed banks c and d that both have the identical string 010000. Furthermore, there appears to be a pattern in these binary strings that distinguishes failed and nonfailed banks. Failed banks tend to have 0 codes and nonfailed banks normally have 1 codes. However, there are some banks that do not have a dominance of either 0 or 1 codes (e.g., banks e and B). Among these banks the positions of the 0 and 1 values in the binary string could be meaningful. Thus, patterns within the codes may be useful in discriminating between failed and nonfailed

banks.

Following this line of reasoning, the string of binary codes is recoded to more fully explore patterns within the binary strings. To do this, a matrix of traits for each bank is created from its binary string. As quoted above in Briggs *et al.*, traits consider all possible combinations of the variables taken one, two, and three at a time and, therefore, seek to find useful interrelations between the variables. Formally stated, each trait (T) is comprised of an array of six integers: T = p, q, r, P, Q, R; where p = 1, 2, ..., L; q = p, p+1, ..., L; r = q, q+1, ..., L; P = 0 or 1; Q = 0 or 1; and R = 0 or 1. The letters p, q, and r act as *pointers* to positions in the binary string from left to right. P, Q, and R give the *values* of the binary code at the positions identified by pointers p, q, and r.[7]

To demonstrate, the trait matrix for failed bank e with binary string 110100 can be developed 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	6	6	100	1	4	5	110
2	2	2	111	3	4	4	011	1	4	6	110
3	3	3	000	3	5	5	000	1	5	6	100
4	4	4	111	3	6	6	000	2	3	4	101
5	5	5	000	4	5	5	100	2	3	5	100
6	6	6	000	4	6	6	100	2	3	6	100
1	2	2	111	5	6	6	000	2	4	5	110
1	3	3	100	1	2	3	110	2	4	6	110
1	4	4	111	1	2	4	111	2	5	6	100
1	5	5	100	1	2	5	110	3	4	5	010
1	6	6	100	1	2	6	110	3	4	6	010
2	3	3	100	1	3	4	101	3	5	6	000
2	4	4	111	1	3	5	100	4	5	6	100
2	5	5	100	1	3	6	100				

As shown, there are 41 traits for this six digit string,[8] which considers all possible

interactions of the variables. It also has a weighting scheme, with both p = q and q = r giving double weight to a particular position (or variable) in the string, and p = q = r giving triple weight to that position. Traits matrices are generated in like fashion for all observations. Of course, as the number of digits in the binary string increases, the size of the traits matrix increases very rapidly due to the factorial mathematics of combinations.

While large traits matrices are possible for problems with only 10 to 20 independent variables, most of the information in each matrix is not useful in distinguishing failing and nonfailing banks and, therefore, can be discarded. In this regard, the traits matrices are trimmed down to include only *features* of nonfailed and failed banks. A feature is a trait that is present relatively frequently in nonfailed (failed) banks but relatively infrequently in failed (nonfailed) banks. For example, if a particular trait (or six digit sequence of binary codes) is found in 80 percent of the failed banks and only 10 percent of the nonfailed banks, it could be defined as a feature of failed banks. We will refer to nonfailed bank features as "good" features and failed bank features as "bad" features. While there is no concrete rule to determine which traits are features, we found that at least 10 to 25 features are needed to obtain good results with the technique. Thus, rules for selecting features that are highly restrictive, as in selecting cutpoints for variables, tend to throw out valuable information that can increase identification accuracy.[9]

After features have been selected, redundant features are dropped. As an example, consider two features (designated 1 and 2) that are found to be present in many of the same nonfailed banks, such that those banks identified correctly by feature 1 are a subset of those banks identified by feature 2. In this instance feature 2 is said to be distinctive (or

dominant), and feature 1 is dropped as being redundant. Elimination of redundant features in this way yields a set of *distinctive features*, each of which provides different information about the observations than the other features. To simplify matters, we eliminated good (bad) features whenever two features were present in the same set of nonfailed (failed) banks.[10]

At this point the good and bad distinctive features can be used to vote on each bank in the sample and then classify the banks. The number of good and bad "votes" (i.e., distinctive features) for each bank are tallied and a voting matrix is constructed. This matrix would have two axes showing the number of good and number of bad votes, respectively, and the number of failed and nonfailed banks in each cell. A cutoff line is chosen from examination of the voting matrix, and observations classified according to their position relative to the cutoff line. A simple rule is to classify banks as nonfailed (failed) when the number of good votes exceeds the number of bad votes. However, this approach assumes that the number of good and bad distinctive features is equal. Since there normally are different numbers of good and bad distinctive features, a cutoff line generally is developed from visual inspection of the voting matrix.

Unlike previous TR studies, we modified the voting procedure by adding an iterative step that enhanced the classification and prediction power of the TR method. Instead of one cutoff line in the voting matrix, two cutoff lines are made in the voting matrix, thus dividing the matrix into three regions. Matrix cells in one region have observations that are only predicted to fail. A second region has matrix cells with only nonfailed bank predictions. The third region, which lies in between the other two regions, contains matrix

cells with mixed predictions of observations, including both failed and nonfailed bank predictions. Observations that do not lie in this middle region of the voting matrix are classified at this point, which we refer to as stage 1. By contrast, observations in the middle region of the voting matrix are used as a separate sample to develop an entirely new TR model (or stage 2 model) with different traits, features, and distinctive features. As before, a voting matrix with two cutoff lines is constructed and observations are classified as before. This iterative procedure is repeated until the number of observations between the two cutoff lines in the voting matrix is approximately constant so that few banks will remain to be classified as failed or nonfailed (i.e., no matrix cells contain only nonfailed or failed banks). We found that the stage 3 model generally had few observations on either side of the cutoff lines, such that we stopped at this third iteration. As we will see in forthcoming results, this iterative method to applying the TR procedure generally improved its ability to classify banks correctly.

Frydman *et al.* have suggested that combining nonparametric and statistical methods could increase classification and prediction power. In the present study we test this proposition by also constructing a hybrid model using both TR and MDA. Instead of using cutoff lines selected by the researcher to classify observations, the hybrid model employs the number of good votes and the number of bad votes as two independent variables in a MDA model. The MDA model provides Z scores with implied probabilities of failure. If a bank had a probability of failure greater than 0.90 (or less than 0.10), it was classified as failed (nonfailed). Observations with Z scores in between this range are used to construct a stage 2 model. This iterative process is continued until no banks have

probabilities of failure (nonfailure) outside the aforementioned range. Again, we found that three stages was sufficient to complete this iterative process in most cases.

Upon completion of this learning phase of the TR program, it can be used to predict bank failure using holdout samples. To gauge the predictive power of the TR results, we compare its performance to MDA and logit models that are commonly found in EWS literature and practice. Ultimately, it is comparative effectiveness relative to established EWS models that is the best measure of TR's performance.

II. Application to Commercial Bank EWSs

Samples and Data

All 145 failed U.S. insured commercial banks in 1986 were gathered, which were the most recent failures at the onset of this research. Failed banks are defined as insured institutions for which disbursements by the Federal Deposit Insurance Corporation (FDIC) were required subsequent to their closure.[11] Financial data were collected for these failed banks both one year and two years prior to failure from year-end Call Report computer tapes in 1984 and 1985, respectively. Due to the unavailability of necessary accounting data (which will be discussed shortly) for some failed banks, the final samples of failed banks in 1984 and 1985 data were 126 and 123 observations, respectively.

Using a random number generator, another sample of 900 nonfailed banks was selected from a population of approximately 15,000 insured U.S. commercial banks on the 1985 Call Report tapes. Unavailability of necessary data for these nonfailed banks in 1984 and 1985 resulted in 878 and 862 observations, respectively. Hence, the total samples for

1984 and 1985 are 1,004 and 985 banks, respectively.

A common discrete choice issue in failure prediction studies is the unequal sampling rate from failed and nonfailed groups. As Maddala (1991) has noted, an adjustment to the constant term can be made to account for this problem when using logit. Also, prior probabilities can be used with MDA and logit to mitigate this potential sampling problem. However, because no prior probability adjustment is made in TR, sample proportions approximating the population would best overcome any sampling bias in the failed/nonfailed observations. Since the failure rate in banking in the sample years was about one percent, but the samples selected suggest a ten percent failure rate, the TR results are at risk (to some extent) of yielding either higher or lower error rates than would be achieved in the banking population (e.g., see Eisenbeis [1977, p. 889]). Even so, the sizes of the nonfailed samples are relatively large compared to previous bank failure studies (e.g., Espahbodi [1991] utilized matched samples of 48 failed banks and 48 nonfailed banks) and, in our opinion, enable a fairly comprehensive evaluation of TR's identification ability. Comparing the distributions of banks by asset size of the nonfailed bank sample and the population of insured U.S. commercial banks, the following frequencies (and proportions in parentheses) result:

	1984		1985	
Bank Size Group	Sample	Population	Sample	Population
Less than $100 million	699 (.80)	12,030 (.83)	675 (.78)	11,767 (.82)
$100 - $500 million	155 (.17)	1,949 (.14)	161 (.19)	2,078 (.15)
$500 - $1,000 million	16 (.02)	203 (.01)	15 (.02)	203 (.01)
More than $1,000 million	8 (.01)	273 (.02)	11 (.01)	309 (.02)
Total	878	14,455	862	14,357

From this descriptive data, the size distribution of the samples of nonfailing U.S. commercial banks in 1984 and 1985 is shown to compare closely with that of the population. Thus, due to the large sample sizes, and the similarity of size distribution to the population, we believe our nonfailed bank samples are representative of the population, at least with respect to size.

Finally, it is important to note that our TR model does not have any sampling advantage relative to the other EWS models tested, and it is this comparative examination that is most relevant in the context of the present study. Since it is well known (e.g., see Espahbodi [1991, p. 66]) that EWS models yield upwardly biased estimates of classification accuracy when employed on the original sample observations, two different holdout sample approaches are used to provide validation tests. First, the 1985 sample data is run through the 1984 model, and the 1984 sample data is run through the 1985 model. This procedure provides information on the stability of the models from year to year. It also allows a realistic application of the EWS models, as their application in practice would require their development in 1984 and subsequent prediction using 1985 data. Second, in both 1984 and 1985 the sample is split -- one-half of the sample is used to develop the model, and the second-half of the sample is run through the model for prediction purposes. This approach gives improved estimates of the accuracy of different models' identification power. Together, these validation tests should enable inferences concerning the relative ability of TR and the hybrid TR/MDA model versus MDA and logit to identify failing banks.

Variables

For each sample bank, 28 financial ratios commonly found in previous bank failure studies (cited earlier) were calculated from the Reports of Income and Condition (Call Reports). Table 1 defines these ratios, which comprise the independent variables. As shown there, the ratio variables proxy a wide variety of financial information, including profit, growth, size, liquidity, loan risk, tax exposure, interest rate risk, loan mix, deposit mix, and capitalization. While the selection of the specific set of variables to measure various aspects of bank condition is somewhat arbitrary, these variables are widely employed by bank regulators, analysts, and academic studies. All ratios are calculated both one year and two years prior to failure. Also, all variables are used in each of the EWS models tested, which is consistent with our purpose of comparing identification accuracy, rather than examining the relative importance of different variables (Eisenbeis [1977]).

Model Comparisons

The results of the pattern recognition techniques are compared to MDA and logit classification models that have been applied to commercial banks as EWSs. Multiple discriminant analysis (MDA) estimates a Z score from a linear model of the following form:

$$Z_i = a + b_1X_{i1} + b_2X_{i2} + \cdots + b_nX_{in}, \qquad (1)$$

where X_{ij} = independent variables, j = 1, ..., n for bank i = 1, ..., m, b_j = the coefficients for the jth independent variable, and Z_i = the linear composite score for the ith bank. We used different prior probabilities of failure (i.e., 0.10, 0.30, 0.50, 0.70, and 0.90) in the SAS

Table 1

Definitions of Financial Ratios

Financial Ratios	Definition
Profit:	
Return on assets	Net income after taxes/Total assets
Return on equity	Net income after taxes/Total equity
Profit margin	Net interest/Total assets
Gross operating margin	(Total operating income - Total operating expenses)/Total Assets
Dividend rate	Cash dividends/Total assets
Size:	
Assets	Total Assets
Growth:	
Capital growth	(Total equity$_t$ - Total equity$_{t-1}$)/ Total equity$_t$
Liquidity:	
Liquid assets	Total security holdings/Total assets
Uninsured deposits	Time deposits more than $100,000/ Total time deposits
Loan Risk:	
Loan exposure	Total loans and leases/Total assets
Loan funding	Total loans and leases/Total deposits
Loan loss rate	Total loan losses/Total loans & leases
Net recovery rate	(Total loan losses - Recoveries)/Total loans and leases
Provision rate	Provision for possible loan losses/ Total assets
Nonaccrual loan rate	(Total nonaccrual loans and leases - Total renegotiated troubled debt)/ Total assets
Past due loan rate	Total loans and leases past due 90 days or more/Total assets
Taxes:	
Tax exposure	Total taxes paid/Total assets
Munis usage	Municipal securities/Total assets

Table 1, continued

Interest rate risk:	
Daily gap	(Adjustable daily assets - Adjustable daily liabilities)/Total assets
Short-term gap	(Assets adjustable 1 day to 1 year - liabilities adjustable 1 day to 1 year)/ Total assets
Long-term gap	(Assets adjustable more than 1 year - Liabilities adjustable more than 1 year)/Total assets
Loan Mix:	
Commercial loan risk	Commercial and industrial loans/Total loans
Real estate loan risk	Total real estate loans/Total loans
Agricultural loan risk	Total agricultural loans/Total loans
Diversification	Sum of squared proportions of the three loan mix ratios for each bank
Deposit mix:	
Demand deposit mix	Demand deposits/Total deposits
Retail deposit usage	Demand and time deposits less than $100,000/Total deposits
Capital:	
Capital ratio	Total equity/Total assets

DISCRIM routine to estimate posterior probabilities of failure (Prob) for each bank. As the prior probability of failure is increased, it is more likely that any given bank will be classified as failed, and vice versa. If Prob < 0.50 (Prob ≥ 0.50), the bank is classified as nonfailed (failed). Misclassifications are recorded for type I errors (i.e., a failed bank classified as nonfailed), type II errors (i.e., a nonfailed bank classified as failed), and total errors.

The logit model estimates the posterior probability of failure for banks as:

$$\log[Prob_i/(1-Prob_i)] = a + b_1X_{i1} + b_2X_{i2} + \cdots + b_nX_{in}. \quad (2)$$

This functional form has some advantages over MDA from a methodological standpoint. For example, MDA requires that the independent variables are multivariate normal, whereas logit assumes a cumulative logistic probability function. According to Espahbodi (1991, p. 56) logit is computationally more tractable than MDA, yields a unique maximum in all cases, and is amenable to alternative nonlinear estimation methods. However, when the assumptions of MDA are met (viz., multivariate normality, equal variance-covariance matrices, and linearity), logit is equivalent to MDA. Thus, MDA is a special case of logit.

III. Empirical Results

In this section the classification and prediction accuracy of TR, TR/MDA, MDA, and logit models are compared. Initially, the discussion focuses on total error rates. While this is a common criteria for EWS performance in past studies, regulators are most concerned with type I errors, in which a failing bank is not predicted to fail. For this reason we also report the results of calculating a weighted efficiencies measure that adjusts identification accuracy for type I error rates.

MDA and Logit Models

Table 2 reports the MDA and logit misclassification results for the original samples. Results are shown for different prior probabilities of failure based on 1984 and 1985 data. As expected, type I errors (missed failed banks) decrease and type II errors (missed nonfailed banks) increase as the prior probability of failure increases. In every case the logit model yields a lower percentage of type I, type II, and total errors than the MDA model. The classification results for the original samples reveal greater percentage misses based on the 1984 data compared to the 1985 data (e.g., in 1984 the minimum misclassification rate is 6.9 percent, while in 1985 the minimum is 3.4 percent). The greater error rate using 1984 data compared to 1985 data also is found in the holdout samples. For example, even though the 1985 logit model has a minimum of 8.1 percent total misses based on 1984 data, running 1985 data through the logit model developed with the 1984 data yields a minimum of 5.4 percent total misses. These results suggest that the ability to predict bank failure is lowered two years prior to failure compared to one year prior to failure. Importantly, the effectiveness of logit from an applied standpoint as an EWS is confirmed by the results for the 1984 model using 1985 data. This model and data yielded a minimum error rate of 5.4 percent, which compares favorably with a minimum error rate of 3.4 percent for the 1985 logit model using 1985 data.

Table 3 gives MDA and logit misclassification results for the split samples. Once again, the logit model is consistently superior to MDA for both 1984 and 1985 data and the different prior probabilities examined. Focusing on the minimum percentage total misses, error rates normally increase in the holdout samples for the MDA and logit models. For

Table 2

MDA and Logit Misclassification Results for Original Samples[1]
(In Percentage)

	Prior Probability of Failure														
	.10			.30			.50			.70			.90		
Model	I	II	Total	I	II	Total	I	II	Total	I	II	Total	I	II	Total
I. Original Samples															
A. MDA															
1. 1984 (n=1,004)	5.5	1.7	7.2	3.8	3.9	7.7	2.9	5.9	8.8	2.1	8.6	10.7	0.7	17.5	18.2
2. 1985 (n=985)	3.6	1.0	4.6	2.6	1.8	4.4	2.1	2.3	4.4	2.1	3.1	5.2	1.2	4.8	6.0
B. Logit															
1. 1984 (n=1,004)	8.4	0.2	8.6	6.6	0.8	7.4	4.6	2.6	7.2	2.7	4.2	6.9	1.0	12.4	13.4
2. 1985 (n=985)	5.6	0.3	5.9	3.8	0.7	4.5	2.5	1.4	3.9	1.3	2.1	3.4	0.7	7.2	7.9
II. Holdout Original Samples															
A. MDA															
1. 1984 Data in 1985 Model	9.7	0.5	10.2	8.7	0.8	9.5	7.1	1.2	8.3	6.5	2.0	8.5	4.7	3.5	8.2
2. 1985 Data in 1984 Model	3.0	2.5	5.5	2.2	5.3	7.5	1.7	7.8	9.5	1.2	12.2	13.4	0.9	20.4	21.3
B. Logit															
1. 1984 Data in 1985 Model	10.0	0.0	10.0	8.5	0.5	9.0	7.5	1.1	8.6	6.1	2.0	8.1	3.3	5.5	8.8
2. 1985 Data in 1984 Model	4.8	1.1	5.9	3.3	2.1	5.4	2.7	4.1	6.8	2.2	6.9	9.1	1.3	14.6	15.9

[1] Type I misclassifications are a failed bank classified as nonfailed, and type II misclassifications are a nonfailed bank classified as failed. In 1984 there are 878 nonfailed banks and 126 failed banks in the sample. In 1985 there are 862 nonfailed banks and 123 failed banks in the sample.

Table 3

MDA and Logit Misclassification Results for Split Samples[a]
(In Percentage)

	Prior Probability of Failure														
	.10			.30			.50			.70			.90		
Model	I	II	Total	I	II	Total	I	II	Total	I	II	Total	I	II	Total
I. Split Samples															
A. MDA															
1. 1984 (n=502)	5.8	2.0	7.8	3.6	3.2	6.8	2.4	5.0	7.4	2.0	7.2	9.2	1.0	15.8	16.8
2. 1985 (n=492)	3.4	1.2	4.6	2.8	2.2	5.0	2.6	2.8	5.4	2.4	4.0	6.4	1.8	6.4	8.2
B. Logit															
1. 1984 (n=502)	8.0	0.2	8.2	6.2	0.6	6.8	3.8	2.2	6.0	2.0	3.8	5.8	1.2	11.0	12.2
2. 1985 (n=492)	6.2	0.2	6.4	3.4	0.6	4.0	2.6	1.4	4.0	1.6	3.0	4.6	0.8	7.4	8.2
II. Holdout Split Samples															
A. MDA															
1. 1984 (n=502)	6.6	1.4	8.0	4.8	3.6	8.4	3.6	4.8	8.4	3.0	7.4	10.4	1.4	16.8	18.2
2. 1985 (n=493)	3.8	0.8	4.6	2.6	1.4	4.0	2.2	2.0	4.2	2.0	3.2	5.2	1.4	5.4	6.8
B. Logit															
1. 1984 (n=502)	11.6	1.8	13.4	5.4	4.0	9.4	3.4	5.6	9.0	1.6	7.2	8.8	0.8	8.8	9.6
2. 1985 (n=493)	8.0	1.6	9.6	2.8	3.0	5.8	2.2	4.4	6.6	1.2	4.6	5.8	0.6	7.2	7.8

[a] Type I misclassifications are a failed bank classified as nonfailed, and type II misclassifications are a nonfailed bank classified as failed. In 1984 there are 878 nonfailed banks and 126 failed banks in the sample. In 1985 there are 862 nonfailed banks and 123 failed banks in the sample.

example, the 1984 (1985) logit model had a minimum error rate of 5.8 (4.0) percent using the original split sample compared to 8.8 (5.8) percent using the holdout split sample. These findings suggest that the EWS models are sample specific to some extent. Nonetheless, the predictive power using holdout samples far exceeds chance, thus recommending their practical application by bank regulators and others.

Trait Recognition Model

Tables 4 and 5 summarize the original sample and split sample misclassification results for different stages of the hybrid TR/MDA model and TR model. Normally, the classification accuracy improves at each iterative stage of the three stage voting procedure described previously; however, in isolated cases an earlier stage model had fewer misclassifications than the stage 3 model. Like the MDA and logit results, minimum error rates are higher for 1984 data relative to 1985 data. For example, referring to Table 4's results for the TR model using the original sample, there is a 1.0 percent minimum error rate with 1984 data versus a 0.1 percent minimum error rate with 1985 data. Also like the MDA and logit results, Table 5 indicates that error rates increase to some extent in the holdout split sample relative to the original split sample.

Comparing the original sample results in Table 4 using TR with those in Table 2 using logit, it is clear that TR outperformed logit in classification ability. In 1984 and 1985 the TR model had minimum total error rates of 1.0 percent and 0.1 percent, respectively, which compares favorably with the minimum total error rates using logit of 6.9 percent and 3.4 percent, respectively. Indeed, the 1985 result is rather remarkable, with 99.9 percent

Table 4

Trait Recognition (TR) Misclassification Results for the Original Sample[1]
(In Percentage)

	Model Stages								
	Stage 1			Stage 2			Stage 3		
Model	I	II	Total	I	II	Total	I	II	Total
I. Original Sample									
A. TR/MDA									
1984 (n=1004)	5.2	4.2	9.4	3.8	2.9	6.7	3.5	2.1	5.6
1985 (n=985)	2.5	2.3	4.9	1.3	1.8	3.1	0.4	1.0	1.4
B. TR									
1984 (n=1004)	1.1	1.2	2.3	0.2	0.8	1.0	0.3	0.9	1.2
1985 (n=985)	0.6	0.4	1.0	0.6	0.3	0.9	0.0	0.1	0.1
II. Holdout Original Sample									
A. TR/MDA									
1984 Data in 1985 Model	6.2	2.6	8.8	5.0	3.5	8.4	8.5	2.2	10.9
1985 Data in 1984 Model	2.3	3.9	6.2	2.3	3.0	5.3	2.4	3.6	6.0
B. TR									
1984 Data in 1985 Model	3.7	0.5	4.2	0.3	1.2	1.5	1.1	0.3	1.4
1985 Data in 1984 Model	0.2	2.8	3.0	0.6	2.4	3.0	1.8	0.9	2.7

[1] Type I misclassifications are a failed bank classified as nonfailed, and type II misclassifications are a nonfailed bank classified as failed. In 1984, there are 878 nonfailed banks and 126 failed banks in the sample. In 1985, there are 862 nonfailed banks and 123 failed banks in the sample.

Table 5

Trait Recognition (TR) Misclassification Results for the Split Sample[1]
(In Percentage)

	Model Stages								
	Stage 1			Stage 2			Stage 3		
Model	I	II	Total	I	II	Total	I	II	Total
I. Original Split Sample									
A. TR/MDA									
1984 (n=502)	6.8	2.0	8.8	3.8	3.2	9.0	3.0	2.8	5.8
1985 (n=492)	3.7	2.6	6.3	2.4	1.8	4.3	1.4	2.4	3.9
B. TR									
1984 (n=502)	0.8	2.6	3.4	0.4	0.6	1.0	0.0	0.2	0.2
1985 (n=492)	1.0	0.4	1.4	0.4	0.6	1.0	0.0	0.2	0.2
II. Holdout Split Sample									
A. TR/MDA									
1984 (n=502)	8.8	1.2	10.0	3.8	4.8	8.6	3.6	4.4	8.0
1985 (n=492)	2.8	2.6	5.5	1.8	1.8	3.7	1.8	2.4	4.3
B. TR									
1984 (n=502)	0.6	3.4	4.0	0.4	1.0	1.4	0.8	1.8	2.6
1985 (n=493)	0.2	0.6	0.8	0.4	0.6	1.0	2.8	1.6	4.5

[1] Type I misclassifications are a failed bank classified as nonfailed, and type II misclassifications are a nonfailed bank classified as failed. In 1984, there are 878 nonfailed banks and 126 failed banks in the sample. In 1985, there are 862 nonfailed banks and 123 failed banks in the sample.

correct classification, and exceeds previously cited bank and business failure studies to the authors' knowledge. The holdout sample prediction accuracy for the TR model declined to 1.4 percent using 1984 data in the 1985 model and 2.7 percent using 1985 data in the 1984 model. Again, this compares favorably to the best logit results of 8.1 percent and 5.4 percent, respectively, representing at least a 50 percent increase in prediction accuracy. Importantly, comparing the results using 1985 data in 1984 EWS models, which is a realistic application of EWS methods, TR's performance of 2.7 percent misses is one-half the 8.1 percent misses by logit.

When using the split sample data, Table 5 shows the minimum misclassification rate for the TR model using both 1984 and 1985 data increased from 0.2 percent in the original split sample to 0.8 percent in the holdout split sample. While this difference is fourfold, it only represents a few observations. More importantly, the minimum misclassification rate of 0.8 percent in the holdout split sample using both 1984 and 1985 data substantially outperforms the logit model, which had a minimum error rate in the holdout split sample of 8.8 percent using 1984 holdout data and 5.8 percent using 1985 holdout data (see Table 3). The *maximum* error rate using TR in the holdout samples for 1984 and 1985 is 4.5 percent, which exceeds the *minimum* error rate in the logit model in comparable tests. In general, the predictive power of the TR model as an EWS is strongly supported when compared to the logit and MDA results.

Interestingly, the TR model almost always did better than the hybrid TR/MDA throughout the iterations. For example, referring to Table 5's results for the original sample using 1984 data, the TR model (without MDA) has a total error rate of 3.4 percent at stage

1 and 0.2 percent at stage 3. By contrast, the TR/MDA model has a total error rate of 8.8 percent at stage 1 and 5.8 percent at stage 3. In Table 3 the logit model for this same year had a minimum total error rate of 5.8 percent, which is comparable to the TR/MDA model. Hence, it does not appears that a hybrid EWS approach improves TR classification or prediction accuracy, as contemplated by Frydman *et al.* Also, the hybrid model does not improve upon using statistical models alone, such as logit.

Comparison of Weighted Efficiencies

Previous discussion reported the total error rates of the different EWS models. Naturally, type I errors involving misclassification of failed banks are more important than type II errors, in which nonfailed banks are misclassified. However, sole comparisons of type I errors assumes that type II errors are irrelevant, thereby giving a skewed view of model performance. A method of focusing attention on missed failed banks that simultaneously takes into consideration total error rates is the weighted efficiency measure (see Korobrow and Stuhr [1985, p. 269]). Following their work, as well as Espahbodi's (1991, p. 67) recent analysis, weighted efficiency measures are calculated as:

$$WE = (FCC/PF) \cdot (FCC/AF) \cdot CC, \qquad (3)$$

where FCC = the number of failed banks correctly classified, PF = the number of banks predicted as failed, AF = the number of banks that actually failed, and CC = the percentage of banks correctly classified. Notice that WE gives the weighted classification score in which the total classification rate is adjusted for correct identification of failed banks. Thus, high WE scores reflect not only a high total classification rate but also a high success

rate in identifying failed banks.

Table 6 shows WEs for the MDA, logit, TR/MDA, and TR models. In every case TR outperforms all other models and, in general, the difference in WE scores is considerable. For instance, using the original sample and 1984 data, TR has a score of 89.88, while MDA, logit, and TR/MDA have scores of 42.45, 51.38, and 55.41, respectively. Using the 1985 data, TR has a WE score of 99.09 in the original sample and 98.19 in the split sample. The holdout WE scores for TR are quite high also -- for example, 87.71 (76.47) when running the 1984 (1985) data through the 1985 (1984) model, and 81.55 (91.23) when running one-half of the 1984 (1985) data through the 1984 (1985) model developed with the other one-half data. Espahbodi (1991, p. 69) reported WE scores in previous banking studies in addition to his study. Scores ranged from a low of 1.54 to a high of 65.00 in other studies, with Espahbodi obtaining a logit score of 58.17 and a MDA score of 50.97, based on 48 failed banks and a matched sample of 48 nonfailed banks in 1983. Thus, while the MDA, logit, and TR/MDA weighted efficiencies shown in Table 6 are similar to those in previous studies, the TR results well exceed norms.

Finally, the appendix provides a discussion of the usage of the independent variables (see Table 1) in the 1984 and 1985 TR models based on the original sample.[12] In brief, all variables were employed by the TR models in at least one of the stages, and many variables are frequently used in the distinctive features of both failed and nonfailed (safe) banks. The frequency of usage of these variables in the distinctive features differed greatly in these two years, which suggests that the *financial profile* of failed banks changes as failure becomes imminent.

Table 6

Comparison of Weighted Efficiencies: MDA, Logit, TR/MDA and TR[a]

Sample and Model		FCC	PF	AF	CC	WE[a]
I. Original Sample						
1984	MDA	71	88	126	92.83	42.45
	Logit	99	141	126	93.13	51.38
	TR/MDA	91	112	126	94.42	55.41
	TR	123	132	126	98.80	89.88
1985	MDA	102	125	123	95.53	64.64
	Logit	98	112	123	96.04	66.95
	TR/MDA	119	129	123	98.58	87.98
	TR	123	124	123	99.90	99.09
Holdout Original Sample						
1984	MDA	79	114	126	91.83	39.90
	Logit	65	85	126	93.13	36.74
	TR/MDA	102	133	126	94.72	59.96
	TR	115	118	126	98.61	87.71
1985	MDA	93	118	123	94.42	56.27
	Logit	90	111	123	94.52	56.08
	TR/MDA	74	108	123	91.57	37.75
	TR	105	114	123	97.26	76.47
II. Split Sample						
1984	MDA	45	61	63	93.22	49.12
	Logit	53	72	63	94.22	58.35
	TR/MDA	48	62	63	94.22	55.58
	TR	63	64	63	99.80	98.24
1985	MDA	44	50	61	95.33	60.51
	Logit	48	55	61	95.93	65.88
	TR/MDA	54	66	61	96.14	69.63
	TR	61	62	61	99.80	98.19
Holdout Split Sample						
1984	MDA	30	37	63	92.03	35.53
	Logit	55	91	63	91.24	48.14
	TR/MDA	44	66	63	92.03	42.85
	TR	59	68	63	97.41	81.55
1985	MDA	49	56	62	95.94	66.35
	Logit	56	79	62	94.12	60.26
	TR/MDA	53	62	62	96.35	70.41
	TR	60	63	62	98.98	91.23

[a]Weighted efficiency (WE) = (FCC/PF)*(FCC/AF)*CC, where FCC = the number of failed banks correctly classified, PF = the number of banks predicted to fail, AF = the number of banks that actually failed, and CC = the percentage of banks correctly classified.

IV. Conclusions and Implications

Record numbers of failures and increased competition in the banking industry has stimulated interest in measuring bank risk among regulatory agencies, investors, and bankers. Research in the field of EWSs has proven that many pending failures can be identified prior to collapse. On-going efforts to refine and improve EWS techniques, such as the present paper, show promise in terms of developing better computer-aided tools for measuring bank risk. At some point, EWS models may well be adapted for application to such regulatory problems as capital adequacy and deposit insurance, both of which are the subject of intense study at the present time, due to the implementation of the new risk-based capital standards effective year-end 1992 and the mandate by the FDIC Improvement Act of 1991 to implement variable-rate deposit insurance by 1994.

This paper has sought to compare the ability of a trait recognition (TR) model and a hybrid TR/MDA model versus MDA and logit models to identify failing banks both one year and two years prior to failure. Previously applied in the geophysical sciences, PT is an heuristic pattern recognition method that is particularly advantageous in identification problems involving many independent variables with unknown interrelationships between one another. A matrix of traits that allows for all possible interactions of the variables taken one, two, and three at a time is unique to the model's construction. Traits that are frequently found in either failing or nonfailing banks, or so-called features, are used as discriminators. Redundant features are dropped to get so-called distinctive features. The last step of this learning phase of the procedure tallies the number of good and bad features for each observation and applies a classification rule to these "votes." In general,

application of the TR model to samples of failed and nonfailed U.S. commercial banks indicates that TR is a viable approach to EWS modeling for banking institutions. As evidenced by the relatively strong performance in terms of classification and prediction analyses and weighted efficiencies scores, TR generally outperformed MDA and logit models, in many cases by a wide margin. Indeed, the identification strength of our TR model, as captured by weighted efficiencies measures, exceeded previously reported findings in related literature. We conclude from these findings that TR is a potentially useful EWS that could aid regulators, bankers, investors, and others in evaluating bank condition.

A major implication of our findings is that there is information contained in interactions between financial ratios that can improve EWS prediction of failing banks. The strong performance of TR relative to MDA and logit can be explained by the fact that TR considers a large number of possible interactions among the independent variables, whereas MDA and logit methods normally ignore these interactions. Of course, it is possible to add interaction variables in MDA and logit models, but as the number of independent variables increases, the large number of possible interactions would diminish (and eventually exhaust) the degrees of freedom. In this case a method of choosing the most useful interaction variables to add to the model would be needed as a first step. Also, if many interaction variables were useful, very large samples would be required for statistical estimation of the model. By contrast, TR is a nonparametric technique that employs computer-intensive search methods for identifying useful traits of failing and nonfailing banks that take into account interactions of independent variables.

Another implication of our research is that further work is recommended to improve TR as a discrimination technique in binary choice problems. Specifically, there is a need to refine methods for choosing cutpoints for binary coding observations on a particular variable, selecting distinctive features from the population of traits, and dividing the voting matrix for classification and prediction of observations. In this regard, it is likely that optimization approaches would enhance TR models, and at the same time, decrease their *ad hoc* reliance on researcher experience and judgement.

Footnotes

1. There is a large body of financial research applying statistical classification models to the identification of failing and nonfailing firms, bond ratings, default and nondefault on commercial loans, and other groupings of observations. Multiple discriminant analysis (Altman, 1968; and Altman, Haldeman, and Narayanan, 1977), regression equations (Edmister, 1972), logit analysis (Ohlson, 1980; and Zavgren, 1983), and probit analysis (Zmijewski, 1984) are examples of these types of models. In general, these techniques generate a score that represents a weighted average of multiple independent variables. Simple cutoff points for the scores are used to determine the grouping of the observations, which is compared to their actual grouping to assess the classification power of the model. For examples of applications to commercial banking, see Stuhr and Van Wicklen (1974), Sinkey (1975), Martin (1975), West (1985), Whalen and Thomson (1988), Gajewski (1990), and Thomson (1991).

2. See Breiman, Friedman, Olshen, and Stone (1984) for detailed discussion of classification and regression trees.

3. See Marais, Patell, and Wolfson (1984) for another application of recursive partitioning (viz., classification of problem loans in commercial banking).

4. For in-depth descriptions of pattern recognition, as well as applications to different disciplines, see Bongard (1970), Tou and Gonzalez (1974), and Niemann (1990).

5. A copyrighted version of the program is available from the authors upon request. A Fortran version for mainframes and a C version for personal computers are available. The C version requires an IBM compatible processor and DOS 3.3 or above operating

Footnotes, continued

system. Professionally written, the PC version has a number of windows that make data input and analysis relatively simple, such that no programming experience is necessary.

6. While no distributional assumptions are required in setting cutpoints for variables, they can be readily introduced. For example, assuming a normal distribution, cutpoints could be set a fixed number of standard deviations from the mean.

7. Previous studies using the Bongard *et al.* algorithm have found that increasing beyond six integers in an array does not significantly enhance the model's classification and prediction power.

8. In general, according to Gelfand *et al.* (1976, p. 229), an observation has $L + C_{L2} + C_{L3}$ traits, where C_{LM} is the number of combinations of L things taken M at a time ($C_{L1} = L$). There are two different traits for fixed p, if P = 0 or 1. There are four different traits for fixed p, q, if $p \neq q$, $q = r$; (P, Q) = (0,0), (0,1), (1,0), or (1,1). There are eight different traits for fixed p, q, r, if $p \neq q \neq r$; (P, Q, R) = (0,0,1) or (0,1,0) etc. up to (1,1,1). Therefore, the total number of different traits possible in strings of length L is: $2L + 4C^2_{L2} + 8C^3_{L3}$.

9. Optimally, it would be most beneficial for the program to automatically conduct a sensitivity analysis of the identification success as features are added. A stopping rule for adding features could then be constructed. The authors are in the process of developing methods to optimize feature selection, as well as the choice of cutpoint

Footnotes, continued

locations. Of course, if the researcher is most concerned about misclassifying failed banks, as opposed to overall identification accuracy of nonfailed and failed banks, the program can offer alternative cutpoints for this purpose.

10. The notion of redundant features lends insight into the coding process of the traits matrix. In the text illustration of the TR procedure, the trait pqrPQR = 166100 is redundant with respect to trait 661001. Both traits would be found (and not found) in the same banks. Thus, the latter trait need not be considered in the trait matrix.

11. FDIC disbursements for closed banks are classified into four categories: deposit payoffs, deposit transfers to operating banks, deposit assumptions, and assistance transactions.

12. The results for the split samples were similar and are excluded for the sake of brevity.

References

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

_____, 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.

Benavidez, A., and M. Caputo (1988), "Pattern Recognition of Earthquake Prone Areas in the Andrean Region," Studie Ricerche, University of Bologna, pp. 133-168.

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.

Bongard, M. (1970), Pattern Recognition, New York, NY: Spartan Books.

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

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

Briggs, P., F. Press, and Sh. 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), "Pattern Recognition of Earthquake Areas in Italy," Physics of the Earth and Planetary Interiors, 21, pp. 305-320.

Edmister, R. (1972), "Financial Ratios and Credit Scoring for Small Business Loans," Journal of Commercial Bank Lending.

Eisenbeis, R. A. (1970), "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.

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.

References, continued

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

Gelfand, I. M., Sh. A. Guberman, M. L. Izvekova, V. I. Heilis Borok, and E. Ranzman (1972), "Criteria of High Seismicity Determined by Pattern Recognition," In Ritzema (ed.), The Upper Mantle, Tectonophysics, 13, 1/4, pp. 415-422.

Gelfand, I., Sh. A. Guberman, V. I. Keilis-Borok, L. Knopoff, F. Press, E. Ya. Ranzman, I. M. Rotwain, and A. M. Sadovsky (1976), "Pattern Recognition Applied to Earthquake Epicenters in California," Physics of the Earth and Planetary Interiors, 11, pp. 227-283.

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

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

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.

Korobrow, L., D. Stuhr, and D. Martin (1976), "A Probabilistic Approach to Early Warning of Changes in Bank Financial Condition," Monthly Review, Federal Reserve Bank of New York, July, pp. 187-194.

Maddala, G. (1991), "A Perspective on the Use of Limited-Dependent and Qualitative Variables Models in Accounting Research," The Accounting Review, 66, pp. 788-807.

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

Niemann, H. (1990), Pattern Analysis and Understanding, Berlin, Germany: Springer-Verlag.

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

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.

References, continued

SAS User's Guide: Statistics (1982), Cary, North Carolina: SAS Institute.

Sinkey, J. E., Jr. (1975), "A Multivariate Statistical Analysis of the Characteristics of Problem Banks," Journal of Finance, 30, pp. 21-36.

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, pp. 233-238.

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

Tou, J. T., and R. C. Gonzalez (1974), Pattern Recognition Principles, Reading, MA: Addison-Wesley Publishing Company.

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, Federal Reserve Bank of Cleveland, 24, pp. 17-26.

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

Zmijewski, M. (1984), "Methodological Issues Related to the Estimation of Financial Distress Prediction Models," Journal of Accounting Research (Supplement), 22, pp. 59-82.

Appendix

Variables Used in TR Models

Tables A-1 and A-2 show the number of times each of the independent variables listed in Table 1 appeared in the distinctive features for the 1984 and 1985 TR models based on original samples. "Safe" distinctive features are used to identify nonfailed banks, while "failed" distinctive features are associated with failed banks.

The most striking feature of the tables is that no variables tended to dominate others in their frequency of use in the TR method. Different models employed in each of the three stages of the TR method had different distinctive features, as evidenced by the changing patterns of variable frequency in these features across model stages. Also, because the variable frequencies considerably differ in Tables A-1 and A-2, the importance of the variables from year to year is not stable (e.g., the provision rate is found frequently in the safe distinctive features in 1984 but is only used in one safe distinctive feature at each stage in 1985). These tabulations of variable usage in different stages and years suggest that the financial profiles of banks changes as the failure becomes imminent.

Another feature of the variable frequencies in Tables A-1 and A-2 is that all variables are employed in each year to classify banks. This differs from statistical techniques in which variables that are highly correlated with other variables are dropped from the model. While two variables may be highly correlated, their interaction with other variables may well differ. By dropping a variable for reasons of high correlation with another variable, information on the interaction of that variable with other variables is lost. The TR technique avoids this potential problem by considering two-variable and three-variable interactions in the development of distinctive features. The usage of all independent variables in both the 1984 and 1985 TR models supports the notion that interactions between variables provides information useful to classification.

Table A-1

Number of Times Variables Appear in Distinctive Features Coding: 1984 Original Sample

	Stage 1		Stage 2		Stage 3	
Variable	Safe	Failed	Safe	Failed	Safe	Failed
Return on assets	11	-	30	45	17	2
Return on equity	3	-	81	17	3	3
Profit margin	-	-	22	24	44	1
Gross operating margin	2	1	48	38	17	5
Dividend rate	17	-	43	5	49	-
Capital growth	7	-	62	29	47	-
Assets	-	5	7	15	1	4
Liquid assets	14	3	50	25	10	-
Uninsured deposits	5	-	48	35	21	9
Loan exposure	10	2	39	34	8	13
Loan funding	5	1	43	13	16	1
Loan loss rate	8	-	8	-	14	6
Net recovery rate	9	-	26	8	16	2
Provision rate	60	2	57	31	28	1
Nonaccrual loan rate	11	-	48	2	36	-
Past due loan rate	5	-	34	8	-	1
Tax exposure	6	1	34	37	2	7
Munis usage	21	-	81	24	21	3
Daily gap	25	-	32	26	54	19
Short-term gap	8	-	89	18	53	1
Long-term gap	20	-	79	10	15	1
Commerical loan risk	5	-	92	14	17	4
Real estate loan risk	-	-	55	23	7	10
Agricultural loan risk	-	-	12	6	29	-
Diversification	-	-	49	20	36	1
Demand deposit mix	3	-	20	6	15	3
Retail deposit usage	4	-	59	3	54	-
Capital ratio	2	-	96	12	11	1

Table A-2

Number of Times Variables Appear in Distinctive Features Coding: 1985 Original Sample

	Stage 1		Stage 2		Stage 3	
Variable	Safe	Failed	Safe	Failed	Safe	Failed
Return on assets	13	3	1	13	-	-
Return on equity	40	-	31	-	1	1
Profit margin	16	-	-	20	-	-
Gross operating margin	18	5	3	12	-	-
Dividend rate	42	13	-	10	1	-
Capital growth	24	2	1	1	1	-
Assets	-	3	5	18	1	4
Liquid assets	15	1	3	10	-	-
Uninsured deposits	13	9	5	9	2	-
Loan exposure	9	1	1	18	2	-
Loan funding	5	6	3	-	2	-
Loan loss rate	-	4	8	6	-	-
Net recovery rate	1	10	1	6	-	-
Provision rate	1	1	1	9	1	-
Nonaccrual loan rate	1	5	9	-	13	-
Past due loan rate	19	-	8	15	-	-
Tax exposure	18	1	33	11	1	1
Munis usage	18	5	10	-	4	-
Daily gap	2	1	12	-	1	-
Short-term gap	18	1	12	11	5	1
Long-term gap	-	5	18	-	1	1
Commerical loan risk	-	1	2	9	4	-
Real estate loan risk	-	-	1	1	1	1
Agricultural loan risk	-	-	12	-	-	-
Diversification	-	2	-	8	-	-
Demand deposit mix	8	-	-	1	1	5
Retail deposit usage	11	1	-	-	3	-
Capital ratio	14	7	13	12	-	1