Using Financial Data to Identify Changes in Bank Condition

By Gary Whalen and James B. Thomson

Gary Whalen is an economic advisor and James B. Thomson is an economist at the Federal Reserve Bank of Cleveland. The authors would like to thank John N. McElravey for research assistance and Ralph L. Day for programming assistance.

Introduction

The 1980s have been characterized by record post-Depression bank failure rates, a record number of banks on the Federal Deposit Insurance Corporation's (FDIC's) problem bank list, and record losses to the FDIC in terms of total dollar losses and losses per dollar of failed bank assets.

Moreover, as the cost and complexity of examining banks have risen it has become increasingly more difficult for the bank regulators to attract and retain quality bank examiners. On the other hand, advances in computer technology give bank regulators the ability to monitor the condition of banks without conducting an on-site examination. Therefore, off-site monitoring of banks has become an important part of the regulatory examination umbrella.

Off-site monitoring tracks the condition of banks using the quarterly call report balance sheet and income statement data.[1] Banking regulators use these early-warning systems to complement on-site examination and as a way to allocate scarce examination resources. When off-site monitoring indicates a deterioration of a bank's financial health, an on-site exam can then be conducted.

The early-warning systems have been developed from an extensive number of studies relating bank condition to bank balance sheet and income statement data. These studies, which use financial data to evaluate financial condition, can be classified into two types. The first type is failed bank studies.[2] These studies use financial data to predict bank failures. Early-warning systems devised from this literature would use the characteristics of failed banks as the benchmark for identifying problem institutions.

The second type of research in this area uses financial data to classify banks into problem and nonproblem categories.[3] In other words, these studies attempt to predict a bank's examination rating using only publicly available data. Our study falls into this class. We use call-report data to predict deterioration in condition as measured by changes in CAMEL ratings.[4] Unlike previous

■ 1 The formal name for the call reports is the Federal Financial Institutions Examination Council's Consolidated Reports of Condition and Income.

■ 2 See Meyer and Pifer [1970], Hanweck [1977], Martin [1977], Pettway and Sinkey [1980], Bovenzi el al. [1983], Rose and Kolari [1985], Wesl [1985], Lane et al. [1986], Sinkey et al. [1987], and Pantalone and Platt [1987].

3 See Stuhr and Van Wicklen [1974], Sinkey [1975, 1977, 1978], Sinkey and Walker [1975], Korobow et al. [1977], and Korobow and Stuhr [1983].

4 CAMEL is an acronym for the five risk categories rated by the bank examiners: Capital adequacy, Asset quality, Management, Earnings, and Liquidily.

studies, however, we are able to include nonperforming loans in the analysis as a measure of asset quality. In addition, we explore the use of factor analysis as a way to statistically mimic the procedure used by examiners to assign CAMEL ratings.

The rest of the paper is organized as follows: Section I reviews the examination process and the assignment of CAMEL ratings as a measure of condition. Section II discusses the role of off-site monitoring in the examination process. Section III describes the data and the basic statistical methods we use in the study. The results of the analysis are reported in section IV and our conclusions appear in section V.

I. The Role of Bank Supervision and Examination in the Regulatory Process

Bank supervision and regulation in the United States is frequently justified by the role that banking plays in the payments system. That is, the safety and soundness of the banking system is perceived to be inexorably intertwined with the stability of the economy. Futhermore, supervision and regulation reduce the moral hazard problem inherent in federal deposit insurance (see Jensen and Meckling [1976], Benston et al. [1986], and Buser et al. [1981]). By identifying problems early, regulators are able to force corrective action, or close the institution in a manner that minimizes losses to depositors and the deposit-insurance fund, and that minimizes the disruptive impact on the economy.

On-site examinations serve four basic functions in the regulatory process. First, they allow bank regulators to determine whether or not the bank is in violation of any state or federal banking laws and regulations. Second, a bank exam may be conducted to evaluate a bank's electronic funds transfer and on-line trading systems. Third, although bank exams are not specifically conducted for the purpose of detecting ill-advised or illegal activites on the part of bank officers, insiders and employees, on-site examinations are an effective method for detecting fraud and malfeasance. In fact, physical inspection of a bank's books is often the only way to detect irregularities in the operation of the bank that may indicate illegal or ill-advised actions by bank employees (see Benston et al. [1986]).[5]

The fourth role of on-site examinations is to determine the financial condition of a bank. Although banks are required to submit quarterly financial statements, known as the Federal Financial Institutions Examination Council's Reports of Condition and Income, to the bank regulators, the best way to determine the quality of a bank's assets and management is still an on-site examination and appraisal of its books and operations.

When the focus of the exam is to determine the financial condition of a bank, the examiner will rate the bank on a scale from one to five (one being the highest) in five basic areas. These five ratings are referred to as CAMEL ratings. The first component of the CAMEL rating is capital adequacy. Bank capital serves as the last line of defense against losses to uninsured depositors, general creditors, and the FDIC. The examiner assesses the level and quality of the bank's capital base and assigns the bank a rating based on that assessment.

Asset quality is the second component of a bank's CAMEL rating. Examiners wade through loan documentation and check the quality of collateral (if any) backing each loan. They make judgements as to the quality of each borrower and his ability to repay the loan. Furthermore, they look for excessive exposure of the bank to a single borrower or industry. The recent problems in the Texas banking industry are a stark reminder of the benefits of portfolio diversification.

The third component of a bank's CAMEL rating is based on the quality of its management. This is the most subjective of the ratings given by the examiner and is often influenced by the quality of the bank's other ratings. The management rating is based on the examiner's perception of the quality of the bank's officers and the efficiency of the management structure.

Earnings is the fourth component of the CAMEL rating. Earnings are rated on both recent performance and the historical stability of the earnings stream. Examiners will look at the composition of bank profits to determine whether they come from a solid operating base or are driven by one-time gains, such as those generated by the sale of assets. Examiners regard earnings as the first line of defense against loan defaults and other unforeseen events.

The fifth component of a bank's CAMEL rating is liquidity. Liquidity is a measure of a bank's

5 Historically, fraud and malfeasance have been a leading cause of bank failures and they still are an important cause of bank failures today. In fact, illegal acts (including fraud, misconduct, and risky speculation) by bank officers, employees, and insiders were cited as the primary cause of failure for over 33 percent of the 138 banks that were closed in 1986 (see Kathleen Doherty, "Who's Minding the Fraud?" *American Banker*, September 21, 1987, p. 15.).

ability to meet unforeseen deposit outflows. This is an important area of risk facing banks because a liquidity crisis may result in the failure of a solvent bank. Examiners look at the bank's funding sources as well as the liquidity of its assets in determining this rating.

The five component ratings are then subjectively weighted by the examiner to arrive at an overall CAMEL rating for the bank. This rating is then used to determine the degree of regulatory attention and resources that will be devoted to the bank. A composite rating of one is thought to indicate a strong bank that could weather adverse economic conditions. A composite rating of two means that the bank could be severely weakened by adverse economic conditions. A three-rated bank is thought to be at risk in an unfavorable economic environment. Four-rated banks are considered to be banks that are in danger of failing unless corrective actions are taken. Finally, a five rating indicates that the bank is likely to fail in the near future.

II. Off-Site Monitoring and Bank Regulation

Although on-site examination of banks is the best tool for determining the financial condition of banks, staff and budget constraints do not allow state and federal banking regulators to examine the majority of banks more frequently than once every 12 to 24 months. The frequency at which a bank is to be examined is determined by its composite CAMEL rating at the time of its last exam. Problem banks (CAMEL rating of three, four, or five) are examined more frequently than banks with composite CAMELs of one or two.

Unfortunately, the condition of a bank may have deteriorated since the time of its last examination and may merit more regulatory scrutiny than its last CAMEL rating indicates. The response to this problem has been the development of off-site monitoring of bank condition or early-warning models using quarterly call report data. Therefore, the off-site monitoring allows more current information to be brought into the supervisory process. When the early-warning system indicates a bank's condition is deteriorating, an exam can be triggered. That is, rather than being a substitute for on-site examination, off-site monitoring is a valuable tool for setting examination priorities. Moreover, because financial conditions tend to deteriorate over time, a reliable early-warning system would allow examiners to devote more time and resources to detecting fraud, malfeasance, and other irregularities in a bank's operations.

Two types of screens have been proposed for use in off-site monitoring. The first type utilizes quarterly balance sheet and income statement data from the call reports. These early warning models construct ratios from the call reports to proxy for the different types of risk targeted in the examination process. For example, published studies of early-warning systems (see Korobow et al. [1977] and Sinkey [1977, 1978]) have used capital-to-asset ratios to proxy capital adequacy. Other ratios such as net charge-offs to total loans, operating income to operating expenses, return on assets, and core deposits to total liabilities are some of the ratios that have been used in these studies to proxy the other four components of the CAMEL rating. Statistical procedures like logit analysis and discriminant analysis are then used to classify banks into problem and nonproblem categories on the basis of the ratios selected.[6]

Sinkey (1977) proposed a second type of early-warning system that uses stock-market data as a screen for deteriorating condition. These models assume stock markets are efficient and that the underlying stochastic process governing stock returns is stable. The market screen for declining condition is based on the analysis of residuals from market model regressions on individual bank stock returns. Tests are performed on these residuals to detect abnormal negative performance by a bank. Negative abnormal performance by a bank's stock indicates a deterioration in its condition. One drawback of this screen is that reliable stock-market data are available only for the largest 100 to 200 banks, making this screen infeasible for the bulk of this country's more than 14,000 banks.[7]

III. Data and Methods

Data Set

The sample of banks analyzed in this study consists of 58 institutions examined by the Supervision and Regulation Department of the Federal Reserve Bank of Cleveland. These banks are located in Ohio, western Pennsylvania, eastern Kentucky, and the panhandle region of West Virginia.

The data set includes at least one actual composite CAMEL rating for each sample bank

☐ 6 Call-report data is also used by bank regulators to construct non-statistical early-warning models that mimic the examination process.

■ 7 A second problem with the stock-market data is that most bank stock is issued at the holding company level. This introduces noise into the market screen.

TABLE 1

List of Variables

Ratio Number	Definition
1	Primary capital/average assets
2	Payout ratio
3	Asset growth rate
4	Net loan and lease charge-offs/average total loans and leases
5	Current recoveries/prior charge-offs
6	Nonperforming loans and leases/primary capital
7	Loans and leases, past-due and nonaccrual/ gross loans and leases
8	Loan loss reserve/total loans and leases
9	Return on average assets
10	Adjusted return on average assets
11	Pretax return on average assets
12	Net interest margin
13	Overhead expense/average earning assets
14	Provision for loan losses/average earning assets
15	Securities gains or losses/average earning assets
16	One year GAP/equity capital
17	One year GAP/total assets
18	Average earning assets/interest bearing liabilities
19	Loans plus securities/total sources of funds
20	Volatile liabilities/total sources of funds
21	Net funds dependency
22	Brokered deposits/total deposits

SOURCE: Authors.

assigned at an on-site examination between November 1983 and July 1986. Several of the banks in the sample were examined more than once over this time period and so a total of 70 composite CAMEL scores were available for the 58 sample banks.

The remainder of the data set is comprised of two sets of financial ratios constructed from publicly available quarterly call-report data. The definition of each ratio used in the study appears in table 1. The financial variables were preselected by the Supervision and Regulation Department of the Cleveland Federal Reserve Bank for use in a nonstatistical early-warning model developed to forecast CAMEL ratings for the same set of sample banks. Thus, each ratio is included because it provides insight on a dimension of the financial condition of the sample banks that is reflected in the actual composite CAMEL rating. The ratios generally are similar to those used in previous early-warning failure-prediction models.

One set of ratios (denoted by the prefix CURR, for current quarter) consists of the ratio values calculated using data from the quarterly call report immediately preceding the date at which the actual composite CAMEL was assigned. If this call date was less than two months before the exam date, the current-quarter ratios were calculated using data from the next closest prior quarter. This was done to reflect the typical two-month lag in the availability of quarterly call data.

The other set of ratios are labeled "previous quarter" (PREV). These are the same set of ratios calculated using call data drawn from reports dated four months before the quarter designated as current.

The Statistical Models

The logit-regression technique was employed to construct several different versions of a model that could be used to predict changes in the CAMEL ratings or, alternatively, the financial condition of the sample banks. Logit analysis was used instead of ordinary least squares or discriminant analysis because the classification accuracy of models estimated using this technique has typically been found to be as good or better than that obtained using other methods.[8]

In all versions of the estimated equations, the dependent variable takes on a value of 1 for sample banks that are categorized as "high risk." These, in turn, are defined to be sample banks with composite CAMEL ratings of 3, 4 or 5. The dependent variable takes on a value of zero for "low risk" banks, in other words, those with CAMEL ratings of 1 or 2.[9]

Two different types of models were then estimated for each set of financial data (that is, "current quarter" and "previous quarter"). In one model, the dependent variable was related to subsets of the ratios appearing in table 1. In the other model, a two-step procedure was

■ 8 For a discussion of logit regression and its relative merits see Bovenzi, et al. (1983), Martin (1977) and Amemiya (1981).

□ 9 The decision to place three-rated banks in the high-risk category is somewhat arbitrary. However, while a CAMEL rating of 3 does not indicate that examiners believe the bank is close to failure, it does reflect their judgment that it is more vulnerable than 1- or 2-rated institutions and that there is need for some corrective action and closer regulatory supervision.

TABLE 2

Logit Model 1 — Large Sample

Variable	Coefficient	T-Statistic	Chi-Square
Constant	-3.48450	-4.61	32.03
CURR06	0.108156	3.40	

	Probability Cutoff Value			
	0.5	0.4	0.3	0.2
Classification accuracy (%)	87.1	88.6	87.1	81.4
Type I error rate (%)	43.8	37.5	31.3	25.0
Type II error rate (%)	3.7	3.7	7.4	16.7

SOURCE: Authors.

TABLE 3

Logit Model 2 — Large Sample

Variable	Coefficient	T-Statistic	Chi-Square
Constant	-4.75058	-1.58	35.51
CURR06	0.093926	2.69	
CURR01	0.101593	0.48	
CURR13	0.355459	0.64	
CURR09	-0.606462	-0.77	

	Probability Cutoff Value			
	0.5	0.4	0.3	0.2
Classification accuracy (%)	90.0	90.0	88.6	85.7
Type I error rate (%)	31.3	31.3	25.0	25.0
Type II error rate (%)	3.7	3.7	7.4	11.1

SOURCE: Authors.

employed. First, factor analysis was used to convert the considerable number of correlated financial ratios into a much smaller number of composite variables or factors that are linear combinations of the original data.[10] The intended result is the creation of a small set of explanatory variables that contains basically the same information as the larger data set. This statistical procedure mimics the procedure used by bank examiners to construct the composite CAMELS assigned at exams. The set of generated factors were then used to construct factor scores for each sample bank. Logit regressions were then estimated using the constructed factor scores as independent variables.[11]

Given the definition of the dependent variable, the estimated coefficients of financial ratios or factors indicative of greater risk or financial weakness (that is, lower capital, lower asset quality, lower earnings, or less liquidity) are expected to be positive.

IV. Empirical Results

Each type of logit model was estimated using three different samples. One, dubbed the "large sample," contained all 70 available observations for the 58 sample banks. Another, labeled the "small sample," contained only one observation for each of the 58 sample banks. These two samples were used to examine the in-sample classification accuracy of the estimated logit models. Since the results using the large and small samples are essentially the same, only the large sample results are reported. The third sample, called the "random sample" is a random sample of 40 banks drawn from the small sample, yielding a hold-out sample of 18 banks. The logit models were then estimated using the sample of 40 banks and used to classify the hold-out sample.

Logit Analysis With Ratio Independent Variables

Estimated logit equations in which subsets of the nontransformed financial ratios were used as independent variables appear in tables 2 to 5. The equations reported are those that did the best job of in-sample classification, using a 50 percent probability cutoff to assign banks to the high-risk group.[12] In-sample classification results are also presented for alternative lower probability cutoff values.

The results demonstrate that the key predictive financial ratio is a measure of asset quality, defined as nonperforming loans and leases

10 The factor-analysis method used is principal-awis factor analysis with prior communality estimates set equal to the squared multiple correlations among variables. The rotation method used was varimax.

11 This is the same approach used in West (1985)

12 The probability cutoff value is the critical value used to assign the sample banks to a risk group, given the prediction of an estimated model. A predicted probability value above the cutoff implies that the bank should be placed in the high-risk group. A cutoff value of 0.5 assumes that the prior probabilities of group membership and the misclassification costs of Type I and Type II errors are equal. Lower cutoff values reflect the view that these assumptions are incorrect.

TABLE 4

Logit Model 3 — Large Sample

Variable	Coefficient	T-Statistic	Chi-Square
Constant	-3.08714	-4.72	27.88
PREV06	0.084124	3.29	

	Probability Cutoff Value			
	0.5	0.4	0.3	0.2
Classification accuracy (%)	87.1	88.6	87.1	80.0
Type I error rate (%)	43.8	37.5	37.5	31.3
Type II error rate (%)	3.7	3.7	5.6	16.7

SOURCE: Authors.

TABLE 5

Logit Model 4 — Large Sample

Variable	Coefficient	T-Statistic	Chi-Square
Constant	-11.98831	-1.67	37.43
PREV06	0.080440	2.35	
PREV01	0.136849	0.64	
PREV13	0.612464	0.92	
PREV09	-2.195222	-2.05	
PREV19	0.082736	1.28	

	Probability Cutoff Value			
	0.5	0.4	0.3	0.2
Classification accuracy (%)	91.4	88.6	85.7	84.3
Type I error rate (%)	31.3	31.3	25.0	18.8
Type II error rate (%)	1.9	5.6	11.1	14.8

SOURCE: Authors.

divided by primary capital (ratio 6). The estimated coefficient on this variable is positive as expected and is statistically significant in almost every case. The results obtained when additional ratios are included are less impressive. The estimated coefficients on the variables are rarely significant and sometimes even exhibit the "wrong sign." Further, adding these variables has only a marginal impact on classification accuracy.[13]

Depending on the sample, model, and chosen probability cutoff value, overall classification accuracy ranges from roughly 82 to 90 percent. For comparative purposes, the classification accuracy of a naive model (which predicts that a bank's current CAMEL is the same as the one assigned at its last exam) is 87.1 percent and 84.5 percent for the large and small samples, respectively.

While the overall classification accuracy of the estimated models is important, judging their usefulness as early-warning tools requires an examination of the Type I (classifying a high-risk bank as a low-risk one) and Type II (classifying a low-risk bank as a high-risk one) error rates of each. Type I errors are typically considered more serious, but if a statistical early-warning model is being developed to aid in the allocation of scarce examination resources, the Type II error rate is also of concern.

Not unexpectedly, the Type I and Type II error rates of the estimated models vary across models and vary with the probability cutoff values used for each one. In general, the Type I error rates are considerable for the estimated models when a 0.5 probability cutoff is employed, while the Type II error rates are very low. The Type I error rates are generally in excess of 30 percent. Reducing the probability cutoff values generally decreases the Type I error rate at the cost of some increase in the Type II rate. When a 0.2 probability cutoff is used (approximately equal to the sample proportion of high-risk banks), the Type I error rate is reduced to roughly 20 percent. The trade-off is a rise in the Type II rate to the 15 percent level. Again, for comparative purposes, the naive model has a Type I error rate of 37.5 percent for the large sample and 46.2 percent for the small one. The Type II error rates are 5.6 percent and 6.7 percent, respectively.

Interestingly, a comparison of the results obtained using current-quarter and previous-quarter ratios indicates only minor differences in the classification accuracy of the estimated models.

Logit Analysis With Factor Scores as Explanatory Variables

Preliminary investigation indicated that most of the variation in the data set could be accounted for by a relatively small number of factors. Accordingly, factor analysis was used on various subsets of the financial ratios to extract two, three or four factors from the sample data. Logit regressions were then estimated using the sets of two-, three-, or four-factor scores produced and used to classify the sample banks into the two risk classes. This exercise revealed that the predictive accuracy of the three-and four-factor

■ 13 This result is similar to Sinkey [1977]. He finds that the ratio of primary capital net of classified assets to total assets (net capital ralio) is Ihe best discriminator between problem and nonproblem banks.

models was no better than that of the two-factor variety. Thus, only the two-factor results are reported and discussed.

The rotated factor-loading matrices for the two-factor models used in the logit regressions reported immediately below appear in tables 6 and 7. These matrices provide insight on the relationship between the observed variables or ratios and the factors produced by the factor analysis. The factor loadings, in turn, are used to generate the coefficients that allow the ratios to be converted into factor scores that are ultimately used as explanatory variables in the logit regressions estimated. Relatively heavy loadings (that is, loadings close to one in absolute value) indicate a close relationship between that variable and the constructed factor and imply that the value of that ratio will have a relatively large impact on the value of the factor score. The sign of the loading indicates the relationship between that particular ratio and the factor score.

In general, an examination of the factor-loading matrices reveals that several asset-quality measures typically cluster together on the first factor. Two other earnings-efficiency-type ratios—return on assets and the overhead expense ratio—also tend to load on factor one, along with the asset-quality ratios. The signs of the loadings on the ratios imply that a sample bank's score on this factor will be higher, the lower its asset quality, the lower its profitability, and the higher its overhead expenses. Thus, higher scores on this factor are indicative of greater risk.

Two liquidity-type ratios—loans plus securities/total sources of funds and volatile liabilities/total sources of funds—typically load together on the second factor. The signs of the loadings imply that scores on this factor will be higher, the higher the former ratio and the lower the latter one. The sign of the loading on the volatile liability ratio suggests that higher levels of this ratio are indicative of more sophisticated liability management and this, in turn, suggests greater liquidity. Higher scores on this factor imply greater liquidity risk.

A third ratio, primary capital/average assets, also tends to load together with the two liquidity ratios. The sign of the loading is positive, implying higher factor scores for banks with higher capital ratios. The reason for the positive loading is unclear.

The estimated logit regressions reported in each table are very similar. In each, the coefficients on the factors exhibit the expected positive signs, but only the coefficient on the asset-quality-earnings factor is statistically significant.

The in-sample classification accuracy of this type of model does not differ markedly from models using simple ratio values. This is true regardless of the sample or type of data employed to construct the factor scores.

When the probability cutoff value is set at 0.5, roughly 90 percent of the sample banks are correctly classified. The Type I error rates of the factor score logits are roughly 30 percent. Type II error rates are generally less than 5 percent. Again, lowering the probability cutoff value lowers the Type I error rate at the cost of an increase in the Type II rate. The Type I error rate remains considerable, hovering around 25 percent even when the probability cutoff value is reduced to 0.2.

As was true for the models in which simple ratios were used, the predictive accuracy of the factor-score models estimated with previous-quarter data is generally no worse and sometimes even slightly better that that of the current-quarter-based counterparts.

TABLE 6

Model 5: Large Sample Rotated Factor-Loading Matrix

CURR06	.886	−.011
CURR07	.872	.071
CURR08	.816	−.018
CURR14	.748	−.050
CURR13	.659	.079
CURR09	−.673	.235
CURR19	.019	.891
CURR01	−.277	.499
CURR20	−.209	−.892

Logit Model 5

Variable	Coefficient	T-Statistic	Chi-Square
Constant	−1.37361	−3.14	37.53
FACTOR1	4.24095	3.31	
FACTOR2	0.86227	0.86	

	Probability Cutoff Value			
	0.5	0.4	0.3	0.2
Classification accuracy (%)	90.0	90.0	85.7	82.9
Type I error rate (%)	31.3	31.3	31.3	25.0
Type II error rate (%)	3.7	3.7	9.3	14.8

SOURCE: Authors.

TABLE 7

Model 6 — Large Sample Rotated Factor-Loading Matrix

PREV06	.909	.010
PREV08	.870	009
PREV07	.866	.040
PREV14	.684	-.124
PREV13	.596	.158
PREV09	-.815	.342
PREV19	.017	.908
PREVO1	-.188	.549
PREV20	-.186	-.880

Logit Model 6

Variable	Coefficient	T-Statistic	Chi-Square
Constant	-1.26098	-3.05	35.10
FACTOR1	4.55155	3.27	
FACTOR2	0.86765	0.93	

	Probability Cutoff Value			
	0.5	0.4	0.3	**0.2**
Classification accuracy (%)	90.0	87.1	85.7	84.3
Type I error rate (%)	31.3	31.3	31.3	18.8
Type II error rate (%)	3.7	7.4	9.3	14.8

SOURCE: Authors.

Out-of-Sample Model Forecasts

Each type of model was reestimated using a randomly selected sample of 40 sample banks and was used to classify a holdout sample of 18 banks. In general, the results mirror the findings already discussed above.

In particular, the most useful current-quarter ratio continues to be the nonperforming-loan ratio. The out-of-sample classification of the estimated equation in which this ratio is the only explanatory variable is relatively accurate, given the small size of the sample being examined. Generally, over 80 percent of the holdout sample is correctly classified. When probability cutoff values of 0.4 and 0.5 are used, the Type II error rate is very low, while the Type I rate is considerable. For lower probability cutoff values, the Type I error rate falls to roughly 20 percent without a marked increase in the Type II rate. When additional ratios are used in the estimated equations, the forecasting performance of the estimated models improves slightly.

The predictive accuracy of the estimated logit models is roughly the same when current-quarter factor scores are used as explanatory variables. This was found to be true regardless of the number of factors employed. The results obtained using previous-quarter data generally mirrored those obtained using current-quarter data.

V. Summary and Conclusion

The results of this study are in accord with those reported by many others who have done previous empirical work on early-warning failure-prediction models. Specifically, the results demonstrate that relatively simple models constructed using only a limited number of financial ratios that are derived solely from publicly available information do a reasonably good job of classifying commercial banks into different risk classes. The overall classification accuracy and Type I and Type II error rates of the models estimated in this study are comparable to those reported by other researchers.[14]

In addition, the critical predictive role of asset quality and earnings measures detected in previous empirical work is confirmed.[15] Particularly noteworthy is the performance of the asset-quality proxy, nonperforming loans divided by primary capital. Models employing only this variable perform as well as more complicated models. Furthermore, nonperforming loans appear to be as good a proxy for asset quality as classified assets derived from examination reports (not publicly available). Previous studies were unable to employ asset-quality proxies using nonperforming loans because it was not available on the call reports before March 1983.

The results actually are somewhat better than expected given a number of circumstances. First, the sample size is very small, much smaller in fact than that used in many previous studies.

■ 14 For example, Wang, et al.(1987) examined a sample of over 2,900 S&L's in a similar study. They report in-sample classification accuracy of 74 percent and Type I and Type II error rates of 31 and 21 percent using a probability cutoff value of 0.5.

■ 15 Asset quality and earnings measures have been found to be significant predictors of bank risk and/or failure in virtually every study reviewed. See, for example, Hirschhorn (1986).

Second, the set of potential explanatory variables was limited at the outset. Given the results obtained in previous work, it is possible that the use of several other variables and/or slightly different versions of ratios actually employed (all of which would be constructed from publicly available data) would have improved the predictive power of the estimated models.

In particular, a size measure might have proven useful, given that the dependent variable is constructed from examiner perceptions of bank risk. It is known that examiners incorporate bank size into their evaluations of the financial condition of banks and a size variable has been found to be useful in previous empirical studies.[16] Loan composition measures such as the ratio of commercial and industrial loans to total loans or assets have been found to be significant predictors of bank risk in other work and may have improved the classification accuracy of the models estimated in this study.[17]

Some researchers have reported that slightly different versions of the ratios available for use in this study improved the predictive power of their models. For example, the ratios of other operating expenses to total assets and primary capital divided by risk assets have been found to be superior to the expense and capital measures used in this study.[18]

Finally, the risk profile of the particular sample of banks used in this study made them difficult to accurately classify with a statistical model. A large proportion (roughly two-thirds) of the sample banks had CAMEL ratings of 2 or 3. Very few of the sample banks had CAMEL ratings of 4 or 5. Thus, the ratio values of the high-risk and low-risk banks in the sample were not markedly different. This may be one reason why the performance of the estimated models was not better and why the results can be characterized as relatively good.[19]

■ **16** A size variable is used in Barth, et al. (1985), Sinkey, et al. (1987), and West (1985), for example. See also the discussion in Bovenzi, et al. (1983), Korobow and Stuhr (1983) and Hirschhom (1986) about the usefulness of size data.

■ **17** The ratio of commercial and industrial loans to total loans was found to be significantly related to bank financial condition in Pantalone and Platt (1987) and Martin (1977), for example.

■ **18** The relative merits of alternative expense measures are discussed in Bovenzi, et al. (1983). The capital-to-risk asset ratio is used in Martin (1977).

■ **19** It should also be noted that the dependent variable is a subjective measure and reflects examiners' perceptions of bank risk. Further, one component of the CAMEL rating that is incompletely reflected in published financial statements is management quality. Thus, an incorrect classification does not necessarily mean that the model is in error.

References

Arnemiya, Takeshi. "Qualitative Response Models: A Survey." *Journal of Economic Literature.* 119 (December 1981): 1483-1536.

Barth, James R., R. Dan Brumbaugh, Jr., Daniel Sauerhaft, and George H. K. Wang. "Thrift Institution Failures: Causes and Policy Issues." *Proceedings From a Conference on Bank Structure and Competition.* Federal Reserve Bank of Chicago (May 1985).

Benston, George J., Robert A. Eisenbeis, Paul M. Horvitz, Edward J. Kane, and George G. Kaufman. *Perspectives on Safe and Sound Banking: Past, Present, and Future.* Cambridge, *MA:* MIT Press, 1986.

Bovenzi, John F., James A. Marino, and Frank E. McFadden. "Commercial Bank Failure Prediction Models." *Economic Review.* Federal Reserve Bank of Atlanta. (November 1983): 27-34.

Buser, Steven A., Andrew H. Chen, and Edward J. Kane. "Federal Deposit Insurance, Regulatory Policy, and Optimal Bank Capital." *Journal of Finance.* 36 (1981): 775-787.

Frane, James W., and Maryann Hill. "Factor Analysis as a Tool for Data Analysis." *Communication in Statistics: Theory and Methods.* A5(6) (1976): 487-506.

Frydman, Halina, Edward I. Altman, and Duen-Li Kao. "Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress." *Journal of Finance.* 40 (March 1985): 269-291.

Hanweck, Gerald A. "Predicting Bank Failure." Working Paper. Board of Governors of the Federal Reserve System, November 1977.

Hirschhorn, Eric. "Developing a Proposal For Risk-Related Deposit Insurance." *Banking and Economic Review.* Federal Deposit Insurance Corporation. (September/October 1986).

. "The Information Content of Bank Examination Ratings." *Banking and Economic Review.* Federal Deposit Insurance Corporation. (July/August 1986).

Jensen, Michael C., and William H. Meckling. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *Journal of Financial Economics.* 3 (1976): 305-360.

Korobow, Leon, and David P. Stuhr. "The Relevance of Peer Groups in Early Warning Analysis." *Economic Review.* Federal Reserve Bank of Atlanta. (November 1983): 27-34.

Korobow, Leon, David P. Stuhr, and Daniel Martin. "A Nationwide Test of Early Warning Research in Banking." *Quarterly Review.* Federal Reserve Bank of New York. (Autumn 1977): 37-52.

Lane, William R., Stephen W. Looney, and James W. Wansley. "An Application of the Cox Proportional Hazards Model to Bank Failure." *Journal of Banking and Finance.* 10 (1986): 511-531.

Lawley, D. N., and A. E. Maxwell. *Factor Analysis as a Statistical Method* New York, NY: American Elsevier Publishing Company, Inc., 1971.

Maddala, G. S. "Econometric Issues in the Empirical Analysis of Thrift Institutions' Insolvency and Failure." Working Paper. Federal Home Loan Bank Board (October 1986).

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

Meyer, Paul A., and Howard W. Pifer. "Prediction of Bank Failures." *Journal of Finance.* 25 (September 1970): 853-868.

Myers, Stewart C. "Determinants of Corporate Borrowing." *Journal of Financial Economics.* 5 (November 1977): 147-75.

Pantalone, Coleen C., and Marjorie B. Platt. "Predicting Commercial Bank Failure Since Deregulation." *New England Economic Review.* Federal Reserve Bank of Boston. (July/August 1987): 37-47.

Pettway, Richard H., and Joseph F. Sinkey, Jr. "Establishing On-Site Bank Examination Priorities: An Early-Warning System Using Accounting and Market Information." *Journal of Finance.* 35 (March 1980): 137-150.

Rose, Peter S., and James W. Kolari. "Early Warning Systems as a Monitoring Device for Bank Condition." *Quarterly Journal of Business and Economics.* 24 (Winter 1985): 43-60.

Rose, Peter S., and William L. Scott. "Risk in Commercial Banking: Evidence from Postwar Failures." *Southern Economic Journal.* 45 (July 1978): 90-106.

Santomero, Anthony M., and Joseph D. Vinso. "Estimating the Probability of Failure for Commercial Banks and the Banking System." *Journal of Banking and Finance.* 1 (1977): 185-205.

Short, Eugenie D., Gerald P. O'Driscoll, Jr., and Franklin D. Berger. "Recent Bank Failures: Determinants and Consequences." *Proceedings of a Conference on Bank Structure and Competition.* Federal Reserve Bank of Chicago, (May 1985): 150-165.

Sinkey, Joseph F. Jr. "Identifying 'Problem' Banks: How do the Banking Authorities Measure a Bank's Risk Exposure?" *Journal of Money, Credit and Banking.* 10 (May 1978): 184-193.

. "Problem and Failed Banks, Bank Examinations and Early-Warning Systems: A Summary." *Financial Crises.* Edward I. Altman and Arnold W. Sametz, Editors. New York, NY: Wiley Interscience Inc., 1977.

______. "A Multivariate Statistical Analysis of the Characteristics of Problem Banks." *Journal of Finance.* 30 (March 1975): 21-36.

______, and David A. Walker. "Problem Banks: Definition, Importance and Identification." *Journal of Bank Research.* (Winter 1975): 209-217.

Sinkey, Joseph F. Jr., Joseph Terza, and Robert Dince. "A Zeta Analysis of Failed Commercial Banks." *Quarterly Journal of Business and Economics.* 26 (Autumn 1987): 35-49.

Spahr, Walter E. "Bank Failures in the United States." *American Economic Review.* 22 (March 1932): 208-238.

Stuhr, David P., and Robert Van Wicklen. "Rating the Financial Condition of Banks: A Statistical Approach to Aid Bank Supervision." *Monthly Review.* Federal Reserve Bank of New York. 56 (September 1974): 233-238.

Wang, George H. K., Daniel Sauerhaft, and Donald Edwards. "Predicting Thrift Institution Examination Ratings." Working Paper 131 (1987). Office of Policy and Economic Research, Federal Home Loan Bank Board, Washington, D.C.

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