

Rating the Financial Condition of Banks: A Statistical Approach to Aid Bank Supervision

By DAVID P. STUHR AND ROBERT VAN WICKLEN*

One of the most important techniques used by bank regulatory authorities in supervising individual commercial banks and evaluating their financial condition is the on-site examination. Over the years, on-site examinations have yielded valuable information on a bank's assets, capital, management, the soundness of its banking practices, and its overall success in serving the community. Such information is used by supervisory personnel at each Reserve Bank to assign a summary rating to the member banks in each Reserve District. The rating is an overall indication of the bank's condition based on the information available from examination reports.

This article reports on an approach that applies statistical techniques for capturing the more important objective and subjective factors that enter the process which Federal regulatory authorities use to examine commercial banks and rate their condition. The project develops a "scoring" technique that provides a measure of the condition of each member bank relative to other member banks in the Second Federal Reserve District. A long-term goal of this project is to identify banking factors that may be used to signal changes in a bank's condition from data available between field examinations.

* David P. Stuhr is an economist in the Banking Studies Department of the Federal Reserve Bank of New York and an Associate Professor of Finance at Rutgers University. Robert Van Wicklen is an assistant economist in the Banking Studies Department. The authors wish to acknowledge the substantial contribution of Leon Korobow, Manager and George R. Juncker, economist in the Banking Studies Department, to the discriminant project and to this article, as well as the valuable programming assistance provided by Christopher Kell of the Data Services Function of the Federal Reserve Bank of New York.

GENERAL SUPERVISORY CRITERIA FOR BANK EXAMINATION RATINGS

In rating the overall condition of a bank, supervisory personnel¹ consider three major factors—i.e., the quality of the bank's assets, the adequacy of its capital, and the caliber of its management. The quality of assets is assessed through a careful analysis of the bank's portfolio during on-site examinations. Those loans, investments, and other assets that, in the judgment of the examiner, involve more than normal risk or have doubtful or loss characteristics are labeled classified assets. Such assets and other loans specially mentioned by the examiners comprise those assets whose quality is below the normal standard of bank assets.

In evaluating a bank's condition, the volume and distribution of those types of assets involving more than normal risk are generally measured in relation to a bank's gross capital funds. The higher the ratio of classified and specially mentioned assets to a bank's gross capital, the greater is the degree of risk to the organization. Capital²

¹ The Federal Reserve has authority to examine all members of the Federal Reserve System but, as a matter of policy and practice, conducts on-site examinations of only state-chartered member banks. The Comptroller of the Currency supervises and examines all national banks. This study is based on data obtained from the reports of examination for both state member banks and national banks in the Second Federal Reserve District.

² Capital includes capital notes and debentures, equity, surplus, undivided profits, reserves for contingencies, and other reserves. Capital notes and debentures, however, represented only a minor portion of the total capital of the banks studied in the years covered by this analysis.

adequacy, therefore, is measured in relation to a bank's ability to absorb losses on its loans or investments as a result of defaults or forced sales at less than original cost.

The caliber of management generally is assessed by supervisory personnel on the basis of the bank's ability to provide safe and competent leadership. An important indication of such leadership is a bank's profitability, but the overall financial condition of the bank also enters into the judgment of examiners in assessing the quality of a bank's management.

THE APPROACH AND THE VARIABLES USED FOR ESTIMATING SUPERVISORY RATINGS

THE SUMMARY RATINGS. In the examination and analysis of each bank, supervisory personnel at each Reserve Bank assign one of four possible numerical summary ratings. They vary from a high of "1" to a low of "4" and depend on an analysis of the quality of the bank's assets, the adequacy of its capital, and the caliber of its management, based on information obtained from the examination report. All banks in category "1" are considered financially strong. This classification encompasses banks that have proven their ability to perform under a wide range of economic and competitive conditions, as well as those banks that have not been fully tested in a competitive environment but whose assets are comprised of a large percentage of loans or investments that entail little or no risk (e.g., United States Government or Government-guaranteed obligations). Banks rated "2" are institutions whose asset quality, capital adequacy, and management capabilities are not quite as strong overall as banks in category "1" but whose financial underpinnings are clearly sound. Banks having summary ratings of "3" and "4" are regarded as weak.

For purposes of this study, banks rated "1" were considered high-rated banks, banks with summary ratings of "3" and "4" were grouped together to form a sample of low-rated banks, and banks with "2" ratings were considered intermediate between the two groups. Using a computer program, a statistical technique known as discriminant analysis then was employed to analyze variables that took systematically different values for high-rated and low-rated banks. The relevant variables were combined into an equation or discriminant function whose weights, or coefficients, computed for each variable maximized the difference between the average score of the high-rated banks and the average score of the low-rated banks, as obtained from the function.

THE EXPLANATORY VARIABLES. In the initial stages of this study, the examiners' primary measure of *asset quality*—

i.e., the ratio of classified and specially mentioned assets to total bank capital—was used. While this measure performed as expected, further investigation revealed that the accuracy of the classifications obtained from the discriminant function could be improved through the use of an alternative measure. This alternative measure was the sum of classified loans, securities, and other assets plus one half of specially mentioned loans, all divided by total loans and securities.

Various measures of *capital adequacy* similar to those calculated by Federal supervisory personnel were employed initially, but generally they did not substantially improve the ability of the function to distinguish between banks with high and low summary ratings. After experimentation with a number of substitute measures of capital adequacy, we found that the ratio of capital to total assets was most effective in enabling the discriminant function to classify the banks correctly according to their respective summary ratings.

The intangible nature of *management quality* required that its influence on the overall summary ratings of commercial banks be indirectly introduced into the discriminant analysis through three proxy variables measuring management performance. A widely known and generally accepted source of such information is the operating ratios published each year by the Federal Reserve. These ratios reflect the ongoing results of management decision making. Two of them—net income before taxes, and dividends, each as a percentage of total capital—contributed to the discriminant function's ability to distinguish between the two groups of banks. In addition to these operating ratios, the ratio of borrowings to total capital was found to aid the discriminant function in capturing aspects of management quality that influence supervisory ratings. In general, the competence of management would be expected to be related positively to the income and dividend variables and negatively to borrowings. However, within limits, a bank's total borrowings may rise in response to stringent credit conditions without any adverse implications for management performance.

Our investigations suggested that *bank size*, as measured by total deposits, contributed to the ability of the function to classify banks according to their summary ratings. Large organizations often are better able to attract competent management and are in a position to diversify their assets and spread portfolio risks.

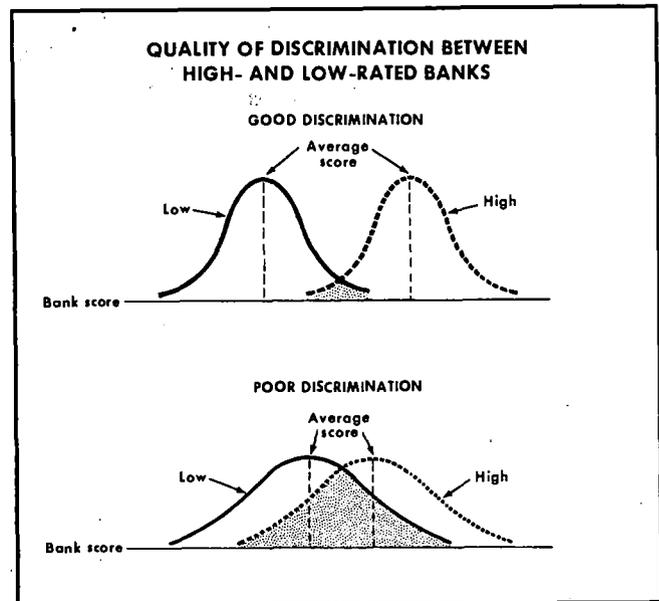
In addition to size, a bank's *organizational structure* is also relevant to bank performance because differences in structure might be expected to result in differences in costs. Given two banks of equal size, one a unit bank and the other having several offices, the latter would be ex-

pected to have a higher cost structure, assuming all other factors are held constant.³ Ideally, the number of banking offices would serve the purpose of capturing differences in cost attributed to organizational structure. However, the required branch data were not compiled as part of the information available for this study and, consequently, the ratio of net occupancy expense to net income was employed as an alternative to the number of offices.

Finally, the *loan-asset ratio* was included to capture differences in ratings that reflected the allocation of a bank's portfolio between relatively higher earning, higher risk loan assets and lower earning, lower risk Government securities and liquidity reserves. Holding all other variables constant, the lower the loan-asset ratio the lower the risk associated with the bank's total assets.⁴

APPLYING DISCRIMINANT ANALYSIS⁵

Discriminant analysis involves the simultaneous study of the effects of a number of variables and, in this study, results in a numerical score for each bank in the sample based on the particular values of the variables enumerated above.⁶ The degree of discrimination between high-rated



and low-rated banks is evaluated by measuring the difference between the average scores of the two groups as well as by how closely the scores are clustered around their respective group averages. The chart illustrates this concept of the quality of discrimination. It indicates that good discrimination occurs when (1) the average scores are widely separated and (2) the individual scores are tightly distributed around their respective group averages. Then the scores of the two groups would have little or no overlap.

Several measures of the quality of the discriminant equation can be calculated. A numerical measure of the likelihood that the equation has successfully divided the sample into two distinct groups is one measure.⁷ Another measure is the likelihood of the equation misclassifying a particular bank; this measure provides an indication of the degree of confidence that may be placed in the function.⁸

³ Other researchers have found that operating costs of banks rise as a result of branching. See, for example, Frederick W. Bell and Neil B. Murphy, *Costs in Commercial Banking: A Quantitative Analysis of Bank Behavior and Its Relation to Bank Regulation*, Research Report No. 41 (Federal Reserve Bank of Boston, 1968) or George J. Benston, "Economies of Scale and Marginal Costs in Banking Operations", *The National Banking Review* (June 1965).

⁴ The data compiled for this study did not include information on the banking markets of individual banks or historical information on the various bank ratios. These types of variables, therefore, were not employed.

⁵ This section is based on material found in J. Johnston, *Econometric Methods* (New York: McGraw Hill, 1972), pages 334-40; C. Kell, "Discriminant Analysis" (Federal Reserve Bank of New York, Research Computer Division, Statistics Section, 1970); G. W. Ladd, "Linear Probability Functions and Discriminant Functions", *Econometrica* (October 1966), pages 873-85; and D. G. Morrison, "On the Interpretation of Discriminant Analysis", *Journal of Marketing Research* (May 1969), pages 156-63.

⁶ Such techniques have been used by other researchers to detect potential weakness or failure. Altman attempted to predict business bankruptcy, while Meyer and Pifer studied bank failures. Dince and Fortson developed a discriminant function to predict bank capital adequacy, which represents only one aspect of the more complex composite rating under study in this article. Cf., Edward I. Altman, "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", *The Journal of Finance* (September 1968), pages 589-609; Robert R. Dince and James C. Fortson, "The Use of Discriminant Analysis to Predict the Capital Adequacy of Commercial Banks", *Journal of Bank Research* (Spring 1972), pages 54-62; Paul A. Meyer and Howard W. Pifer, "Prediction of Bank Failures", *The Journal of Finance* (September 1970), pages 853-68.

⁷ The measure is the F statistic, and it is analogous to the F statistic computed for linear regression analysis. If the F statistic is high, it indicates that there is a significant difference between the group averages.

⁸ Two probabilities for misclassifying an observation can be estimated: (a) the probability of classifying as low a bank whose summary rating was high and (b) the probability of classifying as high a bank whose summary rating was low.

The estimated functions reported below were obtained by including in the data for the high-rated group only the banks that had the highest summary ratings consistently over the entire period from 1964 through 1970. This procedure was used to insure that the data were strongly representative of the characteristics of banks in this group, as defined by the examination process. Because of the small number of banks with low summary ratings in any given year, virtually all of the banks in this category were included in the sample used to estimate the functions. Thus, the sample of banks that was held out was not a randomly selected sample but represented banks that received an intermediate rating at least once during the period 1964-70.⁹

Once the function was computed, a scale adjustment was made so that negative values placed a bank in the low-rated group and positive values placed it in the high-rated group. The scores for the banks in the study then were calculated and ranked in descending order. Scores for those banks that were held out of the original sample were computed by using the weights in the discriminant function, together with the values of the variables in the function. If the function discriminates well, the banks that were accorded high summary ratings should appear near the top of the ranking as a result of having high positive scores, and the banks with low summary ratings should be at the bottom of the ranking as a result of having relatively large negative scores. The banks with intermediate summary ratings should score in between. The position of a bank in the ranking, together with changes in its position over time, should indicate the relative condition of the bank as well as changes in its condition over time.¹⁰ These results were compared with the known rating, thereby providing a check on the accuracy of the function.

Data for this study were available on state-chartered member and national banks in the Second Federal Reserve District. Because of certain limitations, it was not possible to combine all the banks in a single sample. Consequently, four independent functions were estimated: separate ones for state member banks for 1967 and for 1968 and another two for national banks, one for each of these same

⁹ This classification procedure does not strictly comply with the requirement of discriminant analysis, which specifies that each observation be uniquely assigned to a particular group. Program limitations restricted us to the analysis of two groups and necessitated that we exclude intermediate banks in calculating the function.

¹⁰ It should be recognized that such changes in condition cannot be detected by discriminant analysis, unless they become manifest in changes in the variables included in the function.

Table I
SAMPLE BANK CHARACTERISTICS

Category	State member banks		National banks	
	1967	1968	1967	1968
Banks included for estimating function				
High-rated banks	26	26	60	60
Low-rated banks	6	5	18	13
Banks included in holdout group				
High-rated banks	23	22	10	9
Intermediate banks	45	48	22	28
Low-rated banks	1	0	0	0
Total banks in study	101	101	110	263
Banks in Second District	111	108	271	110

years. The sample characteristics of the four functions are given in Table I.

The sample of state-chartered member banks was comprised of all the banks for which complete data were available. Virtually all the national banks with low summary ratings or consistently high ratings were included in the national bank sample. However, many national banks that had at least one intermediate summary rating were omitted from the "holdout" group of banks to keep the sample size manageable.

STATISTICAL PROPERTIES OF THE ESTIMATED FUNCTIONS. The properties of the discriminant functions computed for the years 1967 and 1968 for national and state-chartered member banks are summarized in Table II. The high F statistics indicate significantly different average scores. Probabilities of misclassification range from about 0.1 percent to about 7 percent, indicating only a small overlap in the respective distributions of scores for banks having high or low summary ratings.¹¹

A further check on the quality of the discrimination is possible by observing the pattern of the numbers in Table

¹¹ The signs of the coefficients of the eight variables estimated in the four equations were all correctly predicted by the model, with the exception of two relatively minor variables. The coefficients of the equations are omitted from this article, but a statistical appendix containing these equations will be supplied by the authors on request.

III, which matches the predicted against the actual ratings for all the high- and low-rated banks used in computing the functions. A function can be considered satisfactory if (1) virtually all the banks that were actually accorded high summary ratings also achieved positive discriminant scores from the function with few, if any, receiving negative discriminant scores and (2) virtually all the banks with low summary ratings achieved negative scores from the discriminant function with few, if any, of these banks receiving positive discriminant scores. If discrimination were perfect, the diagonal terms moving from the upper left-hand box to the lower right-hand box in each square would comprise all the observations, while the other two boxes would contain zeros. As can be seen from Table III, only one of the state banks used to compute the function was misclassified—it was one of the five banks that had low summary ratings in 1968; the misclassified bank was accorded a positive discriminant score by the function estimated for 1968. In addition, two banks in the holdout group were misclassified. Both were banks with high summary ratings, but they received negative scores from the discriminant function. During these two years, the functions correctly classified 106 out of 109 state member banks having high or low ratings.

For the sample of national banks, three that had low summary ratings were classified as high by the functions, one in 1967 and two in 1968. Further, in the holdout sample, one bank that had a high summary rating in 1968 was classified as low by the discriminant function. Overall, however, 166 out of the 170 national banks with high or low ratings were classified correctly over the two years. Thus, the fit of the discriminant functions to the process of assigning summary ratings was quite good.

Table II
SUMMARY OF THE PROPERTIES OF THE
FOUR DISCRIMINANT FUNCTIONS

Statistical characteristics	State member banks		National banks	
	1967	1968	1967	1968
F statistic	14.39*	10.23*	20.48*	27.86*
Probability of the function giving a high score to a bank with a low summary rating†	0.60	1.35	7.06	1.84
Probability of the function giving a low score to a bank with a high summary rating†	0.12	0.21	1.62	0.32

* Statistically significant at the 99 percent confidence level.

† In percent.

Table III
ACTUAL RATINGS VERSUS RATINGS OBTAINED
FROM THE DISCRIMINANT FUNCTIONS*

Actual summary ratings	Ratings predicted by the function			
	1967		1968	
	High	Low	High	Low
State member banks				
Banks used in computing the function:				
High	26	0	26	0
Low	0	6	1	4
Banks in the holdout group:				
High	22	1	21	1
Low	0	1	0	0
National banks				
Banks used in computing the function:				
High	60	0	60	0
Low	1	17	2	11
Banks in the holdout group:				
High	10	0	8	1
Low	0	0	0	0

* Evaluates only those banks with high or low summary ratings.

PRELIMINARY IMPLICATIONS OF THE RESULTS

ANALYSIS OF MISCLASSIFICATIONS OF HIGH- AND LOW-RATED BANKS. The potential usefulness of the discriminant function is suggested by an analysis of the ratings of banks that were misclassified by the functions. Two state-chartered member banks in the holdout group—both with high summary ratings—were misclassified by the functions as low performers, one in 1967 and one in 1968. Both banks subsequently received an intermediate rating, indicating that the functions may have been providing some advance indication of a decline in rating. The only state-chartered member bank accorded a low summary rating and misclassified as high by the function in 1968 was barely in the high category of discriminant scores. It was merged out of existence the following year, thus making it impossible to determine whether the summary rating of the bank actually improved.

One national bank with a low summary rating in 1967 achieved a discriminant score which placed it at the bottom of the range that included high-scoring banks. Subsequently, in 1968, that bank was accorded an actual summary rating of "2" (intermediate) which it maintained thereafter. In retrospect, it might appear that the discriminant function was detecting some improvement in the bank's performance.

In contrast, one of two national banks that achieved marginally high scores from the discriminant function, despite low summary ratings in 1968, had been accorded an intermediate summary rating until 1967 and received low summary ratings afterward. In this case the function failed to classify the bank correctly according to either its current or subsequent summary rating. However, the margin of error was small. The other bank had been given a high summary rating for a number of years prior to 1968, was actually accorded a low rating only in 1968, and was upgraded to intermediate in 1969 and 1970. These data suggest that the low summary rating accorded this bank in 1968 may have been due largely to transitory factors. Finally, in the holdout group the one national bank with a high summary rating, but having a low score from the function, subsequently received a low summary rating in 1970. In sum, the functions have provided good discrimination, while a number of apparent misclassifications were suggestive of future changes in the summary ratings.

ANALYSIS OF THE BANKS WITH INTERMEDIATE RATINGS. In the holdout group of banks, thirty-seven state-chartered members had intermediate summary ratings in the years for which the functions were computed and had either negative scores or small positive scores from the discriminant function. These low discriminant scores were suggestive of a tendency toward weakness in the condition of these banks. In reviewing the summary ratings accorded these banks by supervisory personnel, we found that seven of them were given low summary ratings by the end of 1970. (Aside from these seven banks, no other state-

chartered members in the intermediate group received a low summary rating by the end of 1970.) The other thirty banks with actual intermediate ratings and low discriminant scores did not subsequently receive low summary ratings. However, only three of these banks behaved contrary to their function rankings by subsequently obtaining consistently high summary ratings through 1970. The remaining twenty-seven banks were given intermediate summary ratings several times in the period through 1970. A few of these banks had received low summary ratings in prior years, and their low discriminant scores may reflect a borderline status.

Also, in the holdout group, twenty-four national banks had intermediate summary ratings and negative or small positive scores from the discriminant functions. Ten of these banks subsequently received low summary ratings by the end of 1970. In these instances, the discriminant functions provided early indication of the changes. None of the remaining fourteen banks received consistently high summary ratings; six remained consistently intermediate. The eight others had received low ratings in previous years, and perhaps they retained some characteristics that resulted in these low discriminant scores.

CONCLUSIONS

In sum, the results to date are encouraging. The estimated functions do a good job of discriminating between banks having low summary ratings and those whose ratings are high. They also appear to have moderate predictive power. All results must remain tentative, however, until we are able to duplicate the functions' discriminating ability and predictive power over a longer period of time and to establish the stability of the factors that produce accurate discrimination. Moreover, in the longer run, our aim is to develop functions that make use of variables gathered from nonexamination sources so that the early signs of changes in a bank's condition can be available to supervisory personnel in advance of an examination.