FIMS: A New Monitoring System for Banking Institutions

Rebel A. Cole, of the Board's Division of Research and Statistics; Barbara G. Cornyn, of the Board's Division of Banking Supervision and Regulation; and Jeffery W. Gunther, of the Federal Reserve Bank of Dallas's Financial Industry Studies Department, prepared this article.

One of the primary responsibilities of bank regulatory agencies is to minimize the financial loss to the Bank Insurance Fund that results from the failure of insured depository institutions. To discharge this responsibility, bank regulators evaluate the financial performance and condition of depository institutions and initiate prompt corrective actions when they find signs of distress. In the evaluation, regulators use a combination of on-site examinations and off-site monitoring systems.

In 1993, the Federal Reserve instituted the Financial Institutions Monitoring System (FIMS), which is significantly more accurate than previous off-site monitoring systems in identifying financially troubled banking institutions. This article gives the background of FIMS, describes the new system, and explains how it improves on previous systems.

BACKGROUND

As a result of the National Bank Acts of 1863 and 1864, the United States has a dual banking system in which some banks are federally chartered and some are state chartered. The primary bank supervisor and regulator of federally chartered (national) banks is the Office of the Comptroller of the Currency, whereas the responsibility for the supervision and regulation of state-chartered banks is shared by the Federal Reserve, the Federal Deposit Insurance Corporation (FDIC), and the fifty state banking agencies. The primary supervisor and regulator of bank holding companies is the Federal Reserve. Depending upon their activities, bank holding companies may also be subject to regulation by other government agencies, including the Securities and Exchange Commission and the Office of Thrift Supervision.

Under the Federal Deposit Insurance Corporation Improvement Act of 1991 (FDICIA), the bank regulators generally must examine all banks on-site at least once each year; before FDICIA, banks were examined less frequently, except for the state-chartered banks regulated by the Federal Reserve, which were in general subject to annual examinations. FDICIA does not require annual inspections of bank holding companies. According to Federal Reserve policy, bank holding companies that are in sound financial condition are subject to less frequent on-site inspections than are state member banks.

During an on-site examination, regulators visit an institution's offices to evaluate the institution's financial soundness and compliance with laws and regulatory policies, to assess the quality of the institution's management team, and to evaluate the institution's systems of internal control.[1] After the examination, regulators assign the institution a rating that summarizes its financial condition and performance. The rating is known by the acronym

NOTE. This article summarizes the work of a Systemwide Surveillance Task Force composed of economists, examiners, and financial analysts from the Reserve Banks of New York, Philadelphia, Cleveland, Atlanta, Chicago, Minneapolis, Kansas City, Dallas, and San Francisco as well as from the Board of Governors in Washington. We are especially grateful to David Barker and Dale Harrington, who contributed extensively to the project. We also thank David Jones, Myron Kwast, and Sally Davies, who provided comments on earlier drafts.

1. The American Institute of Certified Public Accountants Committee on Working Procedures defines internal control as follows: "Internal control comprises the plan of organization and all of the coordinate methods and measures adopted within a business to safeguard its assets, check the accuracy and reliability of its accounting data, promote operational efficiency, and encourage adherence to subscribed managerial policies."

The Uniform Financial Institutions Rating System

In 1979, federal banking regulatory agencies adopted the Uniform Financial Institutions Rating System as a common way to rate the financial condition of federally insured depository institutions. The system helps identify institutions whose condition warrants special supervisory attention. Under this system, each institution receives a uniform, composite supervisory rating based upon an evaluation of financial performance, condition, operating soundness, and regulatory compliance.

The composite rating of a bank is based upon an on-site evaluation of five critical dimensions of performance—capital adequacy, asset quality, management, earnings, and liquidity; hence the acronym CAMEL, which has become the popularized name for this rating system.

A similar system, known as BOPEC, is used to determine the composite rating for a bank holding company. In the BOPEC system, the composite rating is based upon an evaluation of five elements of the bank holding company—the bank subsidiaries, other (nonbank) subsidiaries, the parent company, consolidated earnings, and consolidated capital adequacy.

In both the CAMEL and BOPEC systems, each component is assigned a rating on a scale of 1 to 5 in descending order of performance:

1—strong performance
2—satisfactory performance
3—performance that is flawed to some degree
4—marginal performance that is significantly below average
5—unsatisfactory performance that is critically deficient and in need of immediate remedial action.

Once the five component ratings have been determined, the composite CAMEL or BOPEC rating is assigned as a summary measure and used by bank regulators as the primary indicator of financial condition.

Composite ratings are assigned on a scale of 1 to 5; 1 indicates that an institution is of least supervisory concern, and 5 indicates that an institution is of most supervisory concern. The five composite rating levels are set forth as follows in the Commercial Bank Examination Manual produced by the Board of Governors of the Federal Reserve System:

1—an institution that is basically sound in every respect
2—an institution that is fundamentally sound, but with modest weaknesses
3—an institution with financial, operational, or compliance weaknesses that give cause for supervisory concern
4—an institution with serious financial weaknesses that could impair future viability
5—an institution with critical financial weaknesses that render the probability of failure extremely high in the near term.

CAMEL, which refers to the five components of the rating system—capital, asset quality, management, earnings, and liquidity (see box "The Uniform Financial Institutions Rating System").

Between on-site examinations, regulators monitor financial institutions off site using computer-based systems. These monitoring systems typically analyze the financial information that each institution must report to regulators quarterly.

Two circumstances in the 1970s prompted the development of such monitoring systems. First, the large number of banking organizations in the United States—more than 14,000 banks and 1,500 bank holding companies as of year-end 1975—and the growing complexity of their financial reports increased the difficulty of systematically analyzing each institution. Second, technological advances in the fields of computer science and data processing significantly reduced the cost of analyzing information. In addition, a precipitous rise in the 1980s in the number of bank failures made clear the need for auxiliary means of supervising banks (see box "The Pattern of Bank Failures since 1980").

Bank Regulatory Surveillance Systems

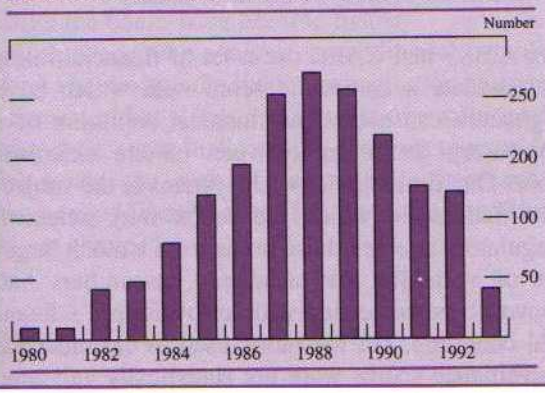
Over the past two decades, various monitoring systems have been developed, but their objectives have generally been the same—to identify developing financial problems at banking institutions between examinations in order to set priorities for the allocation of scarce examination and other supervisory resources. Output from the systems is used to accelerate the on-site examinations of institutions showing financial deterioration; to identify

The Pattern of Bank Failures since 1980

From the mid-1930s until the early 1980s, relatively few banks failed, and losses to the deposit insurance fund were minimal. No more than 20 banks failed in any year. The Depository Institutions Deregulation and Monetary Control Act of 1980 (DIDMCA) set in motion the removal of ceilings on the interest rates that institutions could pay on savings and time accounts and removed or weakened barriers separating commercial banks, thrift institutions, and credit unions. With increased competition, depository institutions weakened by the deep recession of 1981–82 failed at increasingly higher rates (chart). In 1982, 42 banks failed. In each successive year, bank failures rose until 1988, when they peaked at 221. Since then, the number of failures has declined each year; however, it remained in triple digits through 1992, when 122 banks failed. In 1993, bank failures fell to only 41.

From 1982 through 1992, a total of 1,442 banks failed—more than 10 percent of all banks in the United States at the beginning of that period.

Failures of FDIC-insured banks, 1980-93



the areas of most supervisory concern in those institutions scheduled for examination; and to allocate the more experienced examiners to troubled institutions.

Uniform Bank Surveillance Screen

Since the mid-1970s, the Federal Reserve System has monitored the financial performance and condition of banking organizations by screening financial ratios calculated from the Reports of Condition and Income (Call Report) filed quarterly by each banking organization.[2] To improve this monitoring effort, the Federal Reserve System in the mid-1980s adopted the Uniform Bank Surveillance Screen (UBSS) as its primary surveillance system. With some changes, the UBSS remained in service until 1993, when it was replaced by FIMS. The UBSS used financial data from regulatory reports to identify individual institutions whose financial ratios had deteriorated relative to the averages of their respective "peer groups," institutions with similar sizes of assets.[3] The effectiveness of this system, however, was limited by certain methodological weaknesses.

The UBSS was structured around six financial ratios computed from quarterly Call Report data. For both banks and bank holding companies, the first four ratios—tier 1 capital, net income, net liquid assets, and the sum of past due and nonaccrual loans (each expressed as a percentage of total assets)—were the components of a primary surveillance screen.[4] (A surveillance screen uses a set of financial ratio values to identify, or screen, institutions whose condition warrants special supervisory attention.) Within each peer group, the four financial ratios for each institution were sorted from best to worst, and percentile rankings relative to the peer group were calculated. The four ranks were summed to form a bank's composite score, with each rank receiving equal weight in the summation. The resulting composite scores were used to calcu-

2. For a description of the bank surveillance systems used by regulators during the 1970s and early 1980s, see Barron H. Putnam, "Early Warning Systems and Financial Analysis in Bank Monitoring: Concepts of Financial Monitoring," Federal Reserve Bank of Atlanta, *Economic Review* (November 1983), pp. 6–13.

3. The UBSS defined nine peer groups based upon bank asset size: $10 million or less, $10 million–$25 million, $25 million–$50 million, $50 million–$100 million, $100 million–$300 million, $300 million–$1 billion, $1 billion–$3 billion, $3 billion–$10 billion, and greater than $10 billion. A tenth peer group was defined as banks chartered during the previous five years.

4. For banks, asset growth during the previous four quarters and interest paid on volatile liabilities as a percentage of average volatile liabilities were used as supplemental surveillance screens. For bank holding companies, parent company cash flow and double leverage were used as supplemental surveillance screens. As with the four primary ratios, each supplemental screening ratio was converted to a percentile ranking, and institutions with the highest rankings were placed on the exception list for additional off-site analysis and, potentially, for supervisory action. For details on capital standards, see Allan D. Brunner and William B. English, "Profits and Balance Sheet Developments at U.S. Commercial Banks in 1992," *Federal Reserve Bulletin*, vol. 79 (July 1993), pp. 661–62.

late composite percentile rankings within each peer group. These composite percentile rankings served as the basis of the primary surveillance screen. Institutions with the highest composite percentile rankings were placed on an "exception list." Institutions on this list were subjected to more in-depth, off-site analysis by Federal Reserve Bank staff.

The UBSS was supplemented by the quarterly Uniform Bank Performance Report and the Bank Holding Company Performance Report, both from the Federal Financial Institutions Examination Council. These reports are analytical tools created for bank and bank holding company supervisory personnel. In a concise format, they show the effect of management decisions and economic conditions on a banking organization's financial performance and balance sheet composition. The data on performance and balance sheet composition contained in the reports can aid in decisions about capital adequacy, asset quality, earnings, liquidity, and asset and liability management. Each quarterly report shows financial information for multiple time periods. The financial data are presented in the form of ratios, percentages, and dollar amounts. Each report also shows corresponding average data for the institution's peer group along with information identifying how the data ranked the institution relative to its peers.

If this off-site analysis led to the conclusion that the financial condition of an institution had worsened significantly since its most recent on-site examination, a suitable supervisory response was developed and implemented, including contact with the institution's management to obtain additional information and acceleration of the institution's next scheduled on-site examination.

CAEL System

During the mid-1980s, the FDIC developed a surveillance system known as CAEL, which is methodologically similar to the UBSS. The acronym CAEL refers to four CAMEL component ratings that the system evaluates—capital, asset quality, earnings, and liquidity. The system does not provide a management rating. Like the UBSS, CAEL is based upon quarterly bank Call Report data; but whereas the UBSS calculated a composite percentile ranking, CAEL calculates off-site surrogates for CAMEL ratings.

CAEL ratings are calculated in a manner similar to that by which the surveillance scores were calculated in the UBSS, although the calculation of CAEL ratings is considerably more complex and involves many more financial ratios. Like the UBSS, the CAEL system divides banks into peer groups based upon asset size and calculates percentile rankings for four sets of financial ratios that correspond to the four component ratings. Each of the four component ratings is calculated as a weighted average of the corresponding set of financial ratios. The composite CAEL rating is calculated as a weighted average of the four component ratings. Both the ratios used to calculate the ratings and the weights associated with each ratio are determined by a panel of bank examiners. CAEL remains in place today as the FDIC's primary off-site surveillance system.[5]

Limitations of the UBSS and CAEL

The UBSS and CAEL use a set of financial ratios to calculate a composite score with which bank regulators can assess the financial condition of a depository institution between on-site examinations. One limitation of such systems is the subjective manner in which the ratios were selected. Regulators selected these ratios from a much larger set of variables that academic researchers had shown to be correlated with an institution's financial condition, but the ratios used to calculate the surveillance scores were not statistically validated as being sufficiently inclusive to produce accurate off-site assessments of risk. In fact, other ratios, when combined with those of these systems, produce superior assessments of risk.

A related limitation is the manner in which each ratio was weighted. These weights, which were fixed across estimation periods, were determined subjectively rather than by rigorous statistical testing. The UBSS applies equal weights to each of the four financial ratios used to construct the composite surveillance score. CAEL applies a system of weights determined by a panel of senior examiners.

5. The Office of the Comptroller of the Currency relies upon a set of financial ratio screens as its primary surveillance system.

Even if the selected financial ratios contained all the information necessary for an accurate assessment of risk, improper weighting of those ratios would reduce the accuracy of estimation. Moreover, even if optimal weights had initially been assigned, the failure to adjust for temporal shifts would also have reduced estimation accuracy.

A third limitation of these systems is the reliance upon peer-group analysis. Both systems divide banks into peer groups based upon asset size because the average values of key financial ratios are significantly different for banks of different sizes. Without a peer-group analysis, differences in the financial ratios associated solely with bank size could be mistakenly interpreted as differences in financial condition. Because performance is measured relative to that of other banks of similar size, however, systemic changes in the performance either of peer groups or of the banking system as a whole are not incorporated into the composite surveillance scores. Hence, if an entire peer group deteriorates, the percentile scores of individual banks within that peer group may not change, even though the banks have become riskier.

With peer group analysis, an additional complication arises when the size of an institution changes in a manner that places it in a larger or smaller peer group than it was in during the previous quarter. In such a case, the institution's percentile scores may change significantly, even if its financial condition has not changed.

FIMS

Addressing the limitations of the previous off-site bank monitoring systems, FIMS provides two complementary surveillance scores based upon two distinct econometric models—the FIMS rating and the FIMS risk rank. The FIMS rating is an assessment of a bank's current condition, whereas the FIMS risk rank is a longer-term assessment of the bank's expected future condition.

The FIMS rating represents an estimate, based upon the most recent Call Report data, of what a bank's CAMEL rating would be if it were assigned during the current quarter. Because the relationship between financial ratios and CAMEL ratings may change over time, the FIMS rating model is updated each quarter. The updates reflect the most recent relationship between financial ratios derived from the two most recent quarters of bank Call Report data and supervisory ratings based upon the most recent on-site examination. Empirical testing indicates that using data from the two most recent quarters to estimate the historical relationship maximizes the classification accuracy of the rating model.

The FIMS risk rank represents an estimate, based upon a bank's financial condition as measured by the most recent Call Report data, of the probability that a bank will fail during the subsequent two years.[6] Like the FIMS rating model, the risk rank model is updated quarterly to determine which ratios to include and how to weight these ratios. But the risk rank model is updated using financial ratios derived from Call Report data from the same quarter two years previously and information classifying banks as failing or surviving during the intervening period. This procedure enables the risk rank model to incorporate change over time and produces a much longer-term assessment of a bank's financial viability than does the FIMS rating model.

Estimation Techniques

Both the FIMS rating and risk rank are based upon variables representing categories of financial condition. The FIMS rating is based upon the composite CAMEL rating, which can take on integer values from 1 to 5; the FIMS risk rank is based upon a variable that has only two values—0 for failure and 1 for survival. Because such variables represent categories of condition, standard estimation techniques (such as ordinary-least-squares regression analysis) do not provide accurate results. To account for the statistical characteristics of categorical variables, FIMS uses specialized "limited

6. "Failure" is defined as encompassing not only those institutions declared equity insolvent by their primary regulator during the two-year period but also those that are classified as "critically undercapitalized" at the end of the period. The latter group is included to identify institutions for which FDICIA mandates "prompt corrective action." In general, that legislation requires regulators to close critically undercapitalized institutions within ninety days. Critical undercapitalization is defined as a ratio of tangible equity to average assets of less than 2 percent.

dependent variable" estimation techniques to produce its two surveillance scores.[7]

Explanatory Variables

In the literature on financial economics, the numerous studies that model the financial condition of depository institutions show a relatively consistent set of variables to be related to bank financial condition.[8] These variables, which generally include measures of capital adequacy, asset quality, earnings, and liquidity, form the basis for the off-site monitoring systems used by both the Federal Reserve and the FDIC.

To develop FIMS, staff members of the Federal Reserve System selected from the financial literature and the financial ratios commonly used in examination reports approximately thirty financial and structure variables that they considered most likely to be useful in estimating the CAMEL rating and the probability of failure. They tested an additional set of variables measuring regional economic conditions. For the FIMS rating model, the prior-period composite CAMEL rating and the prior-period management component rating were also tested as explanatory variables. The prior-period composite rating was included in the model because the proportion of banks for which the CAMEL rating changes from one examination to the next is less than one-third. The prior-period management component rating was chosen to augment the ability of financial and structure variables to incorporate the management dimension of bank performance into the FIMS rating.

All of the potential explanatory variables except for the prior-period examination rating and the measures of regional economic conditions can be calculated with bank Call Report data (table 1). Income statement variables are based upon data from the previous four quarters. For each variable that is a financial ratio, a four-quarter rate of change is included in the FIMS models. These rates of change are defined as the difference in the values of the current and year-before values of each ratio's numerator, divided by the year-before value of assets.

From this set of potential explanatory variables comes a subset of variables that produces the best estimates of the CAMEL ratings. This subset is selected using a step-wise procedure that evaluates the explanatory power of the entire set of independent variables and sequentially removes from consideration those variables that do not significantly improve estimates of the historical relationship.[9]

A similar procedure is employed in selecting the explanatory variables for estimating the risk rank. From the large set of potential explanatory variables, the subset of variables that produces the best estimate of the probability of failure is chosen. As with the FIMS rating model subset, this subset is selected with a step-wise procedure that first evaluates the explanatory power of the entire set of

7. The ordinal-level logistic regression methodology is used to produce the FIMS rating whereas the binary logistic regression methodology—a special case of the more general ordinal-level methodology—is used to produce the FIMS risk rank.

Each bank receives a set of five estimates representing the probabilities that the next composite examination rating will be equal to 1, 2, 3, 4, or 5. The FIMS rating—or estimated CAMEL rating—is obtained from this set of estimates as the sum of the five possible ratings, each weighted by its estimated probability.

Each bank also receives a single estimate representing the probability that the bank will fail within two years. This estimated failure probability is used to rank banks according to riskiness.

The statistical underpinnings of these methodologies are described in G.S. Maddala, *Limited-Dependent and Qualitative Variables in Econometrics* (Cambridge University Press, 1983). See pp. 22–27 for a description of the binary logistic regression methodology and pp. 46–49 for the ordinal-level logistic regression methodology.

8. For a review of this literature, see Asli Demirguc-Kunt, "Deposit-Institution Failures: A Review of the Empirical Literature," Federal Reserve Bank of Cleveland, *Economic Review*, vol. 25 (Fourth Quarter, 1989), pp. 2–18. Three more recent articles on predicting bank failure are James B. Thomson, "Modeling the Bank Regulator's Closure Option: A Two-Step Logit Regression Approach," *Journal of Financial Services Research* (May 1992), pp. 5–23; Rebel A. Cole and Jeffery W. Gunther, "Separating the Likelihood and Timing of Bank Failure," *Journal of Banking and Finance* (forthcoming); and David S. Jones and Kathy Kuester King, "The Implementation of Prompt Corrective Action: An Assessment," *Journal of Banking and Finance* (forthcoming).

9. To validate this methodology, staff members tested each of the explanatory variables for statistical significance in estimating the CAMEL rating in each quarter from December 1989 through March 1992, a total of ten different estimation periods. Empirical testing verified that inclusion of variables that are not statistically significant often degrades the model's ability to produce accurate estimates for banks not included in the sample used to estimate the model, that is, "out-of-sample" estimates. Empirical testing also has shown that inclusion of variables that consistently are statistically significant improves the ability of FIMS to estimate examination ratings out of sample. In the present context, out-of-sample estimation uses the relationship between the dependent and explanatory variables estimated during one period to estimate events during the subsequent period. Out-of-sample tests comparing predicted and actual outcomes are useful because they most closely resemble the manner in which the model is actually used.

1. Variables for the FIMS rating model

Variable	Description
Loans past due 30–89 days	Loans past due 30–89 days and still accruing interest divided by assets
Loans past due 90 or more days	Loans past due 90 days or more and still accruing interest divided by assets
Nonaccrual loans	Nonaccrual loans divided by assets
Foreclosed real estate	Foreclosed real estate divided by assets
Tangible capital	Equity less goodwill divided by assets
Net income	Net income before extraordinary items less gains or losses on sale of securities divided by assets
Investment securities	Book value of investment securities divided by assets
Reserves	Allowance for loan and lease loss divided by assets
Jumbo CDs	Domestic certificates of deposit (CDs) greater than or equal to $100,000 divided by assets
Net liquid assets	Net liquid assets divided by assets
UBSS asset growth percentile score	Percentile ranking score of growth in total assets over the past four quarters based upon the UBSS system
Volatile liability expense	Interest paid on volatile liabilities divided by average volatile liabilities
UBSS composite percentile score	Composite percentile ranking score based upon the UBSS system
Net charge-offs	Charge-offs less recoveries divided by assets
Brokered deposits	Binary variable equal to 1 if the ratio of brokered deposits to total assets is greater than 1 percent and equal to 0 otherwise
Noninterest expense	Noninterest expense divided by assets
Core deposits	Total deposits less domestic CDs equal to or greater than $100,000, brokered domestic deposits less than $100,000, and foreign deposits divided by assets
Insider loans	Loans to insiders divided by assets
Dividends	Dividends divided by assets
Age	Log of the age of the bank
Size	Log of the current level of assets
Provisions	Provisions for loan and lease loss divided by assets
C&I loans	Commercial and industrial loans divided by assets
Commercial real estate	Commercial real estate loans divided by assets
Consumer loans	Loans to individuals divided by assets
Agricultural loans	Loans for agricultural production divided by assets
Unemployment	Unemployment rate, state level
Income per capita	Personal income divided by labor force, state level
Permits per capita	Housing permits issued divided by labor force, state level

independent variables and then sequentially removes from consideration those variables that do not significantly improve estimates of the probability of failure.[10]

Estimating the Historical Relationship between Call Report Data and CAMEL Ratings

The directions of the estimated historical relationships between the explanatory variables and the CAMEL ratings are shown in table 2. Only variables that are statistically significant in each of the ten quarters are shown. Eleven explanatory variables are statistically significant in each period analyzed. Four of these variables relate to asset quality—the ratios to assets of loans past due 30–89 days and still accruing interest, of loans past due 90 or more days and still accruing interest, of nonaccrual loans, and of foreclosed real estate. Each asset-quality variable is positively related to the numerical CAMEL rating, indicating that higher values of these variables are associated with worse CAMEL ratings.

Of the remaining seven variables that are significant in every period tested, three are negatively

10. This methodology was validated through separate estimations using year-end Call Report data from 1984–88 to determine failures in the two years subsequent to the Call Report date. Out-of-sample estimation accuracy was evaluated for each of these five estimations.

2. Effects of explanatory variables on the FIMS rating

Variable	Effect[1]
Loans past due 30–89 days	Worse
Loans past due 90 or more days	Worse
Nonaccrual loans	Worse
Foreclosed real estate	Worse
Tangible capital	Better
Net income	Better
Investment securities	Better
UBSS asset growth percentile score	Worse
UBSS composite percentile score	Worse
Prior management rating	Worse
Prior composite CAMEL rating	Worse

1. "Worse" indicates that higher values of the variable are associated with worse CAMEL ratings; "better" indicates that higher values of the variable are associated with better CAMEL ratings.

related to the numerical CAMEL rating—the ratios to assets of tangible capital, net income less security gains and losses, and investment securities—indicating that higher values for these variables are associated with better CAMEL ratings. These three variables measure the capital, earnings, and liquidity position of an institution, corresponding to three of the five components of the CAMEL rating system.

The remaining four variables significant in every period are positively related to the numerical CAMEL rating. The UBSS asset-growth and composite percentile rankings are consistently positive, indicating that higher values of these variables are associated with worse CAMEL ratings; also consistently positive are the prior management CAMEL component rating and the prior composite CAMEL rating, indicating that a bank's current rating is a function of its previous ratings. Indeed, a review of the sample banks' ratings reveals that the examination rating of a bank is the same as its previous rating in more than two-thirds of all cases analyzed.

Several additional variables are statistically significant in at least one but no more than five of the ten periods analyzed. Empirical analysis revealed, however, that inclusion of these additional variables in the model does not significantly improve the accuracy of out-of-sample estimation; in most cases, their inclusion usually degrades such accuracy. Of considerable interest is the finding that the regional economic variables tested do not significantly improve out-of-sample estimation. Further analysis indicated that, by themselves, these variables have considerable explanatory power but that this power is attenuated by the inclusion of bank-specific variables in the model.

Estimating the Historical Relationship between Call Report Data and Bank Failure

The directions of the estimated historical relationship between the explanatory variables and the incidence of bank failure are shown in table 3. Only the nine explanatory variables that are statistically significant in each period examined are included in the table. Four of these variables relate to asset quality—the ratios to assets of loans past due 30–89 days and still accruing interest, of loans past due 90 or more days and still accruing interest, of nonaccrual loans, and of foreclosed real estate. Higher levels of each of these variables are associated with a greater likelihood of failure (see note 6 for definition).

3. Effects of explanatory variables on the FIMS risk rank

Variable	Effect[1]
Loans past due 30–89 days	Higher
Loans past due 90 or more days	Higher
Nonaccrual loans	Higher
Foreclosed real estate	Higher
Tangible capital	Lower
Net income	Lower
Reserves	Lower
Investment securities	Lower
Jumbo CDs	Higher

1. "Higher" indicates that higher values of the variable are associated with higher probabilities of failure; "lower" indicates that higher values of the variable are associated with lower probabilities of failure

Of the remaining five variables, four are consistently negative—the ratios to assets of tangible capital, net income, allowance for loan loss, and investment securities—indicating that higher levels of each are associated with a lower likelihood of failure. The coefficient of the final variable—the ratio of domestic certificates of deposit greater than or equal to $100,000 to assets—is positive, indicating that higher levels of this variable are associated with a greater likelihood of failure. This finding is consistent with the financial literature on bank failure, which provides evidence that high-risk banks use volatile liabilities as a funding mechanism to a greater extent than other banks and that these funds can be quickly withdrawn as a bank's condition deteriorates, causing liquidity problems.

Accuracy of Classification

For a surveillance model, the most meaningful measure of accuracy is the ability to classify institutions correctly in a future period rather than the ability to classify institutions correctly in previous periods. Therefore, the following procedure was used to assess the accuracy of the FIMS models. Parameter estimates were generated by applying

the econometric models to Call Report data from the beginning of a given period and to data from events (that is, examinations or bank failures) occurring during that period. These parameter estimates were then applied to Call Report data from the beginning of the subsequent period to generate classification for events occurring during that subsequent period. Finally, these classifications were compared with actual events that occured during the subsequent period.

For example, to assess the accuracy of the FIMS rating model, parameter estimates were generated using data from the March and June Call Report and corresponding examination data from the quarters ending in June and September. These parameter estimates were then applied to September Call Report data to generate estimates of the CAMEL ratings assigned after examinations based upon the September Call Report data.[11] Finally, the estimates based upon the September Call Report data were compared with the actual ratings assigned during examinations based on the same data. This procedure was repeated for ten different estimation periods.

A similar procedure was used to assess the accuracy of the FIMS risk rank model. For example, parameter estimates were generated using data from the December 1984 Call Report and data classifying banks as failing during or surviving through 1985–86. These parameter estimates were then applied to December 1986 Call Report data to classify banks as failing during or surviving through 1987–88. Finally, the classifications based upon the December 1986 data were compared with the actual status of banks at the end of 1988. This procedure was repeated for five different estimation periods.

11. Three official dates are associated with an examination: the date the examination begins, the date the examination ends, and the date of the Call Report data used by the examiners in assigning a supervisory rating. For purposes of estimating and evaluating the accuracy of the FIMS rating model, supervisory ratings are identified by the date of the Call Report data; typically, regulators also identify examinations by that date.

FIMS Rating Model: Estimating the CAMEL Ratings

The broadest measure of estimation accuracy in estimating the CAMEL rating is the ability

4. Accuracy of the FIMS rating and the UBSS composite score in estimating the subsequent-quarter composite CAMEL rating

Actual CAMEL rating	Estimated CAMEL rating										Total
	1		2		3		4		5		
	UBSS	FIMS	UBSS	FIMS	UBSS	FIMS	UBSS	FIMS	UBSS	FIMS	
1											
Number of banks	3,313	4,387	2,295	1,268	48	2	2	0	0	1	5,658
Percentage	58.6	77.5	40.6	22.4	.8	.0	.4	.0	.0	.0	100
2											
Number of banks	2,212	1,339	9,280	11,039	1,970	1,415	343	21	14	5	13,819
Percentage	16.0	9.7	67.2	79.9	14.3	10.2	2.5	.2	.1	.0	100
3											
Number of banks	100	23	1,991	1,188	1,707	2,919	762	503	939	20	4,653
Percentage	2.2	.5	42.8	25.5	36.7	62.7	16.4	10.8	2.0	.4	100
4											
Number of banks	8	0	360	82	732	576	711	1,156	248	245	2,059
Percentage	.4	.0	17.5	4.0	35.6	28.0	34.5	56.1	12.0	11.9	100
5											
Number of banks	2	0	22	6	104	27	255	161	249	438	632
Percentage	.3	.0	3.5	1.0	16.5	4.3	40.4	25.5	39.4	69.3	100
Failed											
Number of banks	1	0	4	0	23	1	73	5	161	256	262
Percentage	.4	.0	1.5	.0	8.8	.4	27.9	1.9	61.5	97.7	100
Total											
Number of banks	5,636	5,749	13,952	13,583	4,584	4,940	2,146	1,846	765	965	27,083
Percentage	20.8	21.2	51.5	50.2	16.9	18.2	7.9	6.8	2.8	3.6	100

NOTE. Based upon bank Call Report data for each quarter from December 1989 through March 1992.

to classify correctly the actual CAMEL ratings of individual banks. To assess the accuracy of the FIMS rating model, one can compare its ratings classifications to those derived from the UBSS—the surveillance system that FIMS replaced. This comparison measures how often each system's estimated quarterly CAMEL rating corresponds with the actual CAMEL rating assigned by examiners based upon the same financial data (table 4).

Table 4 combines classification results from ten separate quarterly estimates based upon the Call Reports from December 1989 through March 1992. For example, the parameters generated from Call Report data for the second and third quarters of 1989 were used to estimate ratings assigned from Call Report data for the fourth quarter of 1989; parameters generated from Call Report data for the third and fourth quarters of 1989 were used to estimate ratings assigned from Call Report data for the first quarter of 1990; and so forth. A total of 27,083 ratings estimates were made.

The FIMS estimates were identical to the subsequently assigned CAMEL rating for 74.6 percent of examinations. Less than 0.5 percent of the estimates were more than one level better than the actual rating, whereas 12.1 percent of the FIMS estimates were exactly one level better than the subsequently assigned CAMEL rating. FIMS was most accurate in estimating CAMEL ratings of 1 (77.5 percent) and 2 (79.9 percent). It also was extremely accurate in identifying banks that failed during the subsequent quarter. Of the 262 failing banks included in the sample, 97.7 percent received a FIMS rating of 5, 1.9 percent received a 4, and the remaining 0.4 percent received a 3; none received a FIMS rating of 1 or 2.

Also in table 4 are the out-of-sample estimation accuracy results for the UBSS. Although the UBSS was not designed specifically to estimate the CAMEL ratings of banks, it did provide a score for each bank, and this score can be used to estimate the examination rating. If all banks are ranked by their UBSS score, CAMEL rating estimates based on the distribution of actual CAMEL ratings can be assigned. For example, if 20 percent of the banks in the sample are 1-rated, 50 percent are 2-rated, 20 percent are 3-rated, 5 percent are 4-rated, and 5 percent are 5-rated, then banks with UBSS scores in the 1st–20th percentiles are classified as 1-rated, and banks with UBSS scores in the 96th–100th percentiles are classified as 5-rated; banks in intervening percentile ranges receive the corresponding ratings.

When ratings estimates were assigned in this manner, the UBSS estimate was equal to the actual CAMEL rating 56.9 percent of the time. Approximately 19.4 percent of the UBSS rating estimates were one level better than the actual CAMEL rating, whereas 2.3 percent of the rating estimates were more than one level better than the actual rating. Like FIMS, the UBSS was most accurate in estimating CAMEL ratings of 1 (58.6 percent) and 2 (67.2 percent), but these percentages were much lower than those for FIMS (77.5 percent and 79.9 percent, respectively.) The UBSS also was much less accurate than FIMS in identifying banks that failed during the subsequent quarter. Of the 262 failing banks in the UBSS sample, only 61.5 percent received a 5-rating; 27.9 percent received a 4-rating; 8.8 percent received a 3-rating; and 1.9 percent received a 1- or 2-rating.

FIMS Rating Model: Identifying Unsatisfactory Banks

Regulators often divide banks into two broad groups—those that are satisfactory and those that are unsatisfactory. In defining satisfactory banks, regulators typically label banks with CAMEL ratings of 1 or 2 as satisfactory and banks with ratings of 3, 4, and 5 as unsatisfactory. As a second measure of estimation accuracy, this classification scheme was used to analyze the ability of the FIMS rating model and the UBSS to classify banks correctly as satisfactory or unsatisfactory.

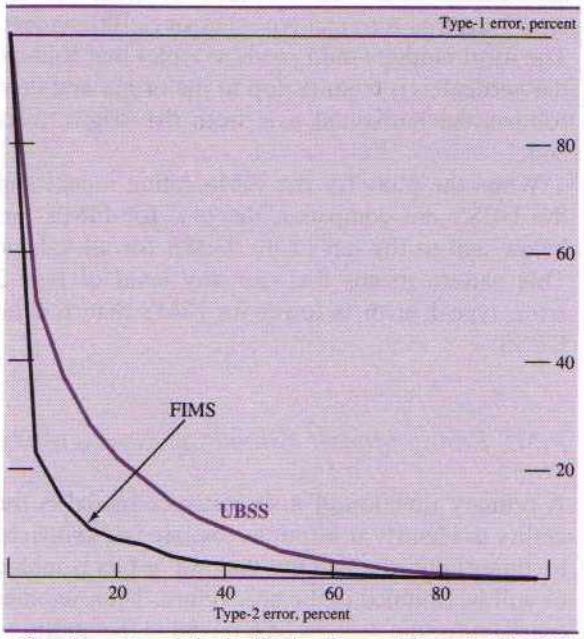
Two types of errors can be made in using an off-site monitoring system to classify banks in this manner. First, banks that actually are unsatisfactory can be misclassified by the system as satisfactory. Misclassification of unsatisfactory banks as satisfactory is referred to as a "type-1 error." The second type of error is to misclassify satisfactory banks as unsatisfactory, a "type-2 error." The cost of a type-1 error can be high because it can result in a bank failure that might have been prevented by early supervisory intervention. The cost of a type-2 error is usually much lower because it is limited to the sum of the unnecessary expenditure of supervisory or examination resources on a healthy bank

and the costs of examination that are borne by the bank.

The accuracy of the FIMS rating model and of the UBSS in identifying satisfactory and unsatisfactory banks is compared in table 5. FIMS incorrectly identified approximately 17.1 percent of the unsatisfactory banks as satisfactory (type-1 error) while incorrectly identifying 7.4 percent of the satisfactory banks as unsatisfactory (type-2 error). The UBSS incorrectly identified approximately 32.7 percent of the unsatisfactory banks as satisfactory (type-1 error) and incorrectly identified 12.2 percent of the satisfactory banks as unsatisfactory (type-2 error).

The information in table 5 is based on the assumption that a FIMS rating model score of 2.5 differentiates satisfactory banks (scores of 2.5 or less) from unsatisfactory banks (scores greater than 2.5). The ability of the FIMS rating model and the UBSS to identify unsatisfactory banks can be increased by adjusting the cutoff score between satisfactory and unsatisfactory downward from 2.5. For example, FIMS scores of 2.3 or less could be classified as satisfactory, whereas scores greater than 2.3 could be classified as unsatisfactory. Such an adjustment would increase the number of banks classified correctly as unsatisfactory and decrease type-1 error, but at the cost of decreasing the number of satisfactory banks correctly classified and increasing type-2 error. A larger percentage of the unsatisfactory banks would be identified, but a larger percentage of satisfactory banks would be misclassified.

1. Comparison of type-1 and type-2 error rates in distinguishing satisfactory and unsatisfactory banks through FIMS and UBSS



NOTE. A type-1 error is the classification of an unsatisfactory bank (CAMEL 3, 4, or 5) as satisfactory; a type-2 error is the classification of a satisfactory bank (CAMEL 1 or 2) as unsatisfactory.

Chart 1 demonstrates this trade-off graphically for the FIMS rating model and for the UBSS. Each line in the figure starts at the upper left corner because labeling no banks as unsatisfactory implies that all of the truly unsatisfactory banks are mislabeled, so that type-1 error is 100 percent and type-2 error is zero. Similarly, each line in chart 1

5. Ability of the FIMS rating and UBSS composite score to correctly identify satisfactory and unsatisfactory banks

Actual CAMEL rating	Estimated CAMEL rating				Total
	Unsatisfactory		Satisfactory		
	UBSS	FIMS	UBSS	FIMS	
Unsatisfactory					
Number	5,118	6,307	2,488	1,299	7,606
Percentage	67.3	82.9	32.7	17.1	100
Satisfactory					
Number	2,377	1,444	17,100	18,033	19,477
Percentage	12.2	7.4	87.8	92.6	100
Total					
Number	7,495	7,751	19,588	19,332	27,083
Percentage	28.6	28.6	71.4	71.4	100

NOTE. A "satisfactory bank" is a bank with a composite CAMEL rating of 1 or 2, whereas an "unsatisfactory bank" is a bank with a composite CAMEL rating of 3, 4, or 5. These results are based upon FIMS ratings calculated using bank Call Report data from each quarter from December 1989 through March 1992 and upon the composite CAMEL ratings assigned in each subsequent quarter.

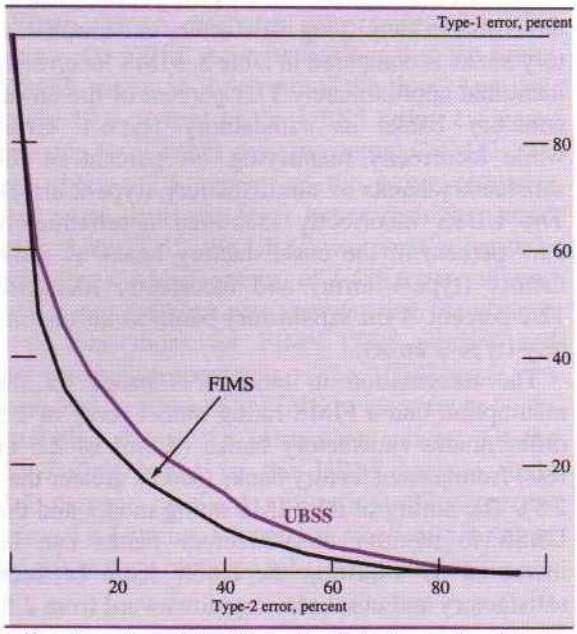
ends at the lower right corner, because labeling all banks as unsatisfactory implies that none of the satisfactory banks are correctly labeled, so that type-1 error is zero and type-2 error is 100 percent. The ideal model would produce a plot that follows the vertical axis from its top to the origin and then follows the horizontal axis from the origin to its end.

When the plots for the FIMS rating model and the UBSS are compared, the plot for FIMS lies below and to the left of the UBSS for all values. This pattern means that, for any level of type-2 error, type-1 error is lower for FIMS than for the UBSS.

FIMS Rating Model: Estimating Downgrades

A primary function of a surveillance model is the ability to identify institutions that are not known to be financially troubled but that are in fact troubled or will be troubled in the near future. Thus, another criterion for the success of a model is the ability to identify those banks that are rated satisfactory (CAMEL 1 or 2) but that will be downgraded to unsatisfactory (CAMEL 3, 4, or 5) in the near term. Once again, a trade-off exists between type-1 and type-2 error rates (table 6). In this case, a type-1 error occurs when the model incorrectly classifies a downgraded bank, and a type-2 error occurs when the model misclassifies a bank that is not downgraded.

2. Comparison of type-1 and type-2 error rates in identifying the downgrading of banks through FIMS and UBSS



NOTE. A type-1 error is the failure to identify the downgrading of a bank; a type-2 error is the false identification of a downgrading.

According to table 6, FIMS incorrectly labeled 58.8 percent of downgraded banks as satisfactory (type-1 error) and incorrectly labeled only 2.7 percent of the CAMEL 1- or 2-rated banks as a downgrade (type-2 error). By comparison, the UBSS incorrectly labeled 55.5 percent of downgraded

6. Ability of the FIMS rating and the UBSS composite score to identify banks downgraded from a composite CAMEL rating of 1 or 2 to a composite CAMEL rating of 3, 4, or 5

Actual CAMEL rating	Estimated CAMEL rating				Total
	Unsatisfactory		Satisfactory		
	UBSS	FIMS	UBSS	FIMS	
1 or 2					
Number	2,002	481	16,092	17,613	18,094
Percentage	11.1	2.7	88.9	97.3	100
3, 4, or 5 (downgrade)					
Number	1,019	757	818	1,080	1,837
Percentage	45.5	41.2	55.5	58.8	100
Total					
Number	3,021	1,238	16,910	18,693	19,931
Percentage	15.2	6.2	84.8	93.8	100

NOTE. Based upon FIMS ratings calculated with bank Call Report data from each quarter from December 1989 through March 1992 and upon the composite CAMEL ratings assigned in each subsequent quarter. Sample includes only banks that had previously been rated as satisfactory; therefore total is smaller than that in tables 4 and 5.

banks as satisfactory (type-1 error) and incorrectly labeled 11.1 percent of the CAMEL 1- or 2-rated banks as downgrades (type-2 error). Hence, the type-1 error rate for the UBSS is slightly less than that of FIMS, but the type-2 error rate is much greater than that of FIMS.

As with the distinction between satisfactory and unsatisfactory tested earlier, comparing the type-1 versus type-2 error trade-off over all possible cut-off values is a more revealing test. Such a comparison of FIMS and the UBSS in identifying downgraded banks versus satisfactory banks appears in chart 2. As before, the plot for FIMS lies below and to the left of that for the UBSS for all values, demonstrating that, for any level of type-2 error, type-1 error is lower for FIMS than for the UBSS.

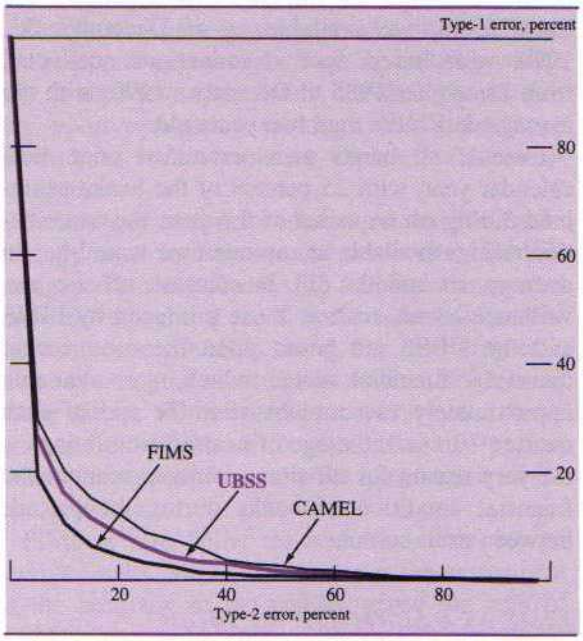
FIMS Risk Rank Model: Accuracy in Estimating Bank Failures

To assess the accuracy of the FIMS risk rank model in estimating the likelihood of bank failures, out-of-sample estimates of the probability of failure within a two-year period were calculated using binary logistic regression methodology. The accuracy of out-of-sample estimation was assessed over the five two-year periods beginning with year-ends 1986–90. For comparison, estimates of failure over these same two-year periods were constructed for the UBSS by ranking banks from worst to best based upon their UBSS composite percentile scores. Altogether, 48,306 estimates were made over the five periods.

As an additional test of accuracy, banks were ranked by their CAMEL rating as of year-ends 1988, 1989, and 1990 to see how well the CAMEL rating estimated failures during the subsequent two-year period relative to FIMS and the UBSS. Over these periods, a total of 32,306 estimates were made using each system.

The type-1 and type-2 error rates for each system were calculated and are plotted in chart 3. In this chart, the vertical axis represents the proportion of failing banks incorrectly identified as surviving (type-1 error), and the horizontal axis represents the proportion of surviving banks incorrectly identified as failing (type-2 error). The lines plotted on these axes represent the trade-off between these two types of error.

3. Comparison of type-1 and type-2 error rates in identifying bank failures through FIMS, UBSS, and CAMEL



NOTE. A type-1 error is the failure to identify the failure of a bank; a type-2 error is the false identification of a bank failure.

The classification accuracy for each of the three models is good, as indicated by the high degree of curvature in the plots. The plots demonstrate that the FIMS rating model is more accurate than the UBSS or CAMEL, as the FIMS curve lies below and to the left of the UBSS and CAMEL curves. For example, when 5 percent of the surviving banks are misclassified, FIMS misclassifies 20 percent of the failing banks. In comparison, the UBSS misclassifies 28 percent of the failing banks and CAMEL misclassifies 32 percent of the failing banks. When 10 percent of the surviving banks are misclassified, FIMS misclassifies 9 percent of the failing banks; the UBSS, 16 percent; and CAMEL, 22 percent. With the current population of approximately 11,000 banks, to reduce the percentage of misclassified failing banks to 9 percent the UBSS and CAMEL would have to misclassify approximately 800 more and 1,300 more surviving banks as failed, respectively, than would FIMS.

The relatively poor performance of the CAMEL rating is most probably attributable to the fact that CAMEL ratings available at any given date are based upon information that is more dated than that

for the off-site monitoring systems. In many cases, these examinations occurred more than a year before the date of interest. For example, the CAMEL ratings available as of December 31, 1990, were based upon examinations conducted from December 1985 to December 1990, with the average data more than two years old.

Even if all banks were examined once each calendar year, with 25 percent of the banks examined during each quarter of the year, the examination ratings available at any one time would be, on average, six months old. In contrast, off-site surveillance scores such as those produced by FIMS and the UBSS are based upon the most recent quarterly financial data, which are available approximately two months after the end of each quarter.[12] In fact, the age of examination ratings is the very reason for off-site systems to monitor the financial condition of banks during the periods between examinations.

FIMS as a Surveillance Model for Bank Holding Companies

As part of its regulatory responsibilities, the Federal Reserve is responsible for supervising bank holding companies. The Federal Reserve uses the so-called BOPEC system for rating the financial condition of bank holding companies as determined from on-site inspections. A BOPEC rating consists of a composite rating derived from five component ratings plus a separate management rating. The five component ratings are for the "bank," "other," "parent," "earnings," and "capital" components (hence the acronym BOPEC). The first three components refer to the three segments of the consolidated bank holding company—its bank subsidiaries covered by the Bank Insurance Fund, its other subsidiaries, and its parent company. As with the CAMEL rating, each component rating and the composite rating are scaled from 1 to 5. The separate management rating, however, has only three levels—"S" for satisfactory, "F" for fair, and "U" for unsatisfactory. Thus, a bank holding company receiving the highest possible ratings would have a BOPEC of "11111/1–S."

12. For expositional purposes, this discussion assumes that only one date is associated with each examination when, in actuality, three are. See note 11.

FIMS provides the Federal Reserve with a means for estimating the bank component of the BOPEC rating. Because the bank component rating is very highly correlated with the composite BOPEC rating, this estimate can serve as an off-site surveillance rating for bank holding companies. For a multibank holding company, the FIMS rating is calculated as the asset-weighted average of its subsidiary banks' FIMS ratings. For a one-bank holding company, it is the same as the subsidiary bank's rating.

FIMS also provides a risk rank for the combined bank portion of bank holding companies. Like the FIMS rating for bank holding companies, the FIMS risk rank for a multibank holding company is calculated as the asset-weighted average of its subsidiary banks' FIMS risk ranks, whereas for a unitary bank holding company it is the same as the subsidiary bank's risk rank. Because bank assets comprise the vast majority of a bank holding company's consolidated assets, these asset-weighted risk-rank averages should provide a fairly reliable off-site assessment of a bank holding company's financial condition when used in conjunction with off-site monitoring of the nonbank subsidiaries and consolidated organization.

Conclusion

The Financial Institutions Monitoring System has been developed to provide the Federal Reserve System with estimates of the financial condition of commercial banks and savings banks insured by the Bank Insurance Fund between on-site examinations. FIMS has several advantages over the Federal Reserve's previous off-site surveillance systems and the expert-based models used by other federal regulators.

First and most important, the accuracy of the new system in estimating the financial condition of banks as indicated both by subsequent on-site examination ratings and by subsequent failures is superior to that of the Federal Reserve's previous model.

Second, the new system provides objective measures of a bank's financial condition. Both the variables and the variable weights that are used to calculate these measures are determined by rigorous statistical testing rather than by subjective judgment.

Third, the new system provides a consistent measure of banks' financial condition. Both models that make up the new system can be calculated for each bank.

Fourth, the new system provides a timely measure of financial condition. The FIMS rating and risk rank for an individual bank can be calculated as soon as the bank files its quarterly Call Report rather than later, when enough quarterly Call Report data are available to calculate meaningful peer-group averages.

Fifth, the new system is more flexible than alternative systems. Explanatory variables can be added to or deleted from FIMS with minimal revisions to software or procedures. The UBSS and CAEL use fixed sets of financial ratios to calculate the surveillance scores, and any change in these ratios would require considerable revision to the surveillance system. The greater flexibility of FIMS should enable staff members at the Board and the Reserve Banks to continue to improve the new system's accuracy over time as experimentation with different variables continues and as feedback from end-users is incorporated into the system. Moreover, because the coefficients on the explanatory variables change each quarter in reflection of the changing conditions in the banking industry, FIMS should continue to be more accurate than existing alternative systems.

Finally, the new system can identify deterioration or improvement in the banking industry within peer groups and systemwide. Unlike systems that rely upon peer-group rankings, FIMS measures absolute as well as relative changes in financial condition.

Preliminary testing has indicated that the methodology used to estimate the composite CAMEL rating produces estimates of the five component CAMEL ratings that are as accurate as estimates of the composite CAMEL rating. By providing estimates of component ratings as well as of the composite rating, FIMS could be used to better focus examination efforts on the dimensions of performance that appear to require the most urgent supervisory attention.

FIMS is also being tested for possible use on foreign banks. Most problematic is the assessment of the accuracy of the results, given the lack of CAMEL ratings for foreign banks. Comparison of FIMS ratings for foreign banks with alternative measures of risk, however, suggest that the FIMS approach is a promising avenue of research. □