Chapter 13

Off-Site Surveillance Systems

Introduction

Bank regulators have used computerized off-site surveillance systems since 1975, yet during the banking crises of the 1980s and 1990s, bank supervisors seemed surprised as each new bank crisis erupted. This chapter examines why, even with computerized off-site systems, it is difficult to anticipate which banks will fail many years in advance of the failure and what tools bank regulators can use to identify banks in the various stages of financial distress.

A brief history of off-site monitoring and a discussion of the advantages and disadvantages of off-site systems are followed by a section in which the goals of a forecasting system are discussed, two sample approaches to achieving those goals are explored, and the conclusion is drawn that the best way to predict long-term failure rates is to measure risk characteristics. The next section focuses on the obstacles to predicting failures in the real world (the life cycle of failing banks, the role of the economic environment, and the non-linear nature of banks' financial process), and the following section develops and tests an analysis of risk groups. Then the systems currently in use at the three bank supervisory agencies are described, with special attention to the FDIC's systems for monitoring growth and tracking changes in bank financial condition that may warrant added supervisory attention. (Also included is a discussion of several proposed improvements in the FDIC's Growth Monitoring System.) A brief concluding section sums up the lessons learned, given the history of banking in the 1980s and early 1990s and the strengths and limitations of current computerized off-site surveillance systems.

History of Off-Site Surveillance Systems

The advent of computerized off-site monitoring of banks in 1975 significantly affected bank examination and enforcement in the 1980s and 1990s. Computerized systems allowed regulators to analyze rapidly and systematically the enormous amounts of data that banks report on their Call Reports. Back in the 1960s, when computers and computer time were very expensive, there were no off-site monitoring systems as we understand them today. But from the early 1960s onward the price of computer time kept dropping,[1] and dramatic price drops in the early 1970s coincided with a crisis at the Office of the Comptroller of the Currency (OCC). Two large national banks failed, United States National Bank (USNB) in 1973 and Franklin National Bank in 1974.

In response to the USNB failure, the OCC commissioned a study by the accounting firm of Haskins & Sells to recommend changes in the OCC's examination system. The report, issued in 1975, recommended putting less reliance on comprehensive reviews of assets in the OCC's banks, increasing the reporting by banks, and establishing a computerized off-site system.[2] It also recommended making vast changes in examination procedures, and implementation in 1976 resulted in a sharp drop in the annual number of on-site examinations, mainly by extending the time between examinations from 12 months to 18 months.[3] In 1975 the OCC did institute an off-site system, the National Bank Surveillance System, in which the primary tool was the Bank Performance Report (BPR).[4] The surveillance system drew on early economic research into the causes of bank failure and on the OCC's own analysis, and the BPR used various financial ratios and benchmarks of financial performance for different "peer groups" to identify banks that could develop problems.[5]

The Haskins & Sells recommendations were designed to make the OCC examination system more efficient, but in the early 1980s the computerized ability to analyze Call Report data was used to help justify reducing the frequency of on-site bank examinations and therefore the number of bank examiners.[6] In fact, between 1975 and 1983 the OCC became

---

[1] For example, the System/360 Model 30 IBM mainframe computer, introduced in 1964, had a price-per-instruction-per-second cost of $25.02 in 1992 dollars. In 1971 IBM released the System/370 Model 135, after which the price per instruction dropped to $8.91 (1992 dollars), a 65 percent decrease. Throughout the 1970s the rate of price decreases accelerated. In 1979 IBM released the 4341 mainframe, with which the price per instruction fell to $0.64 (1992 dollars), less than 10 percent of the 1971 price (Emerson W. Pugh, *Building IBM: Shaping an Industry and Its Technology* [1995], 329).

[2] Eugene N. White, *The Comptroller and the Transformation of American Banking, 1960-1990* (1992), 27, 38–39.

[3] In 1976 there were 5,426 examinations in 4,737 national banks; in 1977 there were 2,886 examinations in 5,665 national banks, a 47 percent decline (White, *Comptroller*, 38). For a more detailed description of these changes, see Chapter 12.

[4] Before this time the Call Report itself was the principal off-site monitoring tool. Examiners would look at their particular institution's Call Report to see if there were any significant changes from the previous examination or the previous Call Report.

[5] Edward I. Altman, "Predicting Performance in the Savings and Loan Association Industry," *Journal of Monetary Economics* 3 (October 1977): 443–66; and Joseph Sinkey, "A Multivariate Statistical Analysis of the Characteristics of Problem Banks," *Journal of Finance* 30 (March 1975): 21–36. The National Bank Surveillance System eventually became the Uniform Bank Surveillance System (UBSS), and the Bank Performance Report became the Uniform Bank Performance Report (UBPR). Currently (1997), the UBPR is the major tool used by banks and bank regulators to compare an individual bank's performance with the performance of its peers. Information on peer groups is given below, in the section entitled "Modifying the Peer Groups."

[6] White, *Comptroller,* 61; and Linda W. McCormick, "Comptroller Begins Major Revamp," *American Banker* 147 (April 29, 1982), 15. See Chapter 12.

so identified with computerized off-site monitoring that the cake at the OCC's 120th-anniversary celebration was in the shape of a computer.[7]

During this same period the Federal Reserve Board (FRB) and the FDIC developed their own off-site systems similar to the OCC's.[8] However, as the number of bank failures dramatically increased through the early 1980s, it became obvious that off-site monitoring was not a substitute for frequent, periodic on-site examinations but was instead a valuable complement to the examination process and could be used to target examination resources. Examinations provide a scrutiny of management practices that no Call Report can capture, and makes it possible for loans to be reviewed in detail. Moreover, studies have shown that examinations affect the integrity of Call Reporting by encouraging banks to recognize loan losses in a timely manner. And unless Call Report data are accurate, an off-site system will not be effective.[9]

To make surveillance systems more useful, changes were introduced in the early 1990s. As a result, contemporary bank surveillance systems are designed to take Call Report data and build indicators of the condition of a bank so that regulators can determine whether additional supervisory attention is warranted before the next regularly scheduled on-site examination. Regulators have also developed various failure models that predict how many banks have a high probability of failure within the next two years. These models are used to plan for the FDIC's future cash needs and to alert examiners to the impending failures.

Advantages and Disadvantages of Off-Site Monitoring

The best way for supervisors to track the condition of banks is to conduct frequent, periodic on-site examinations of banks. But examiners cannot be perpetually on-site at all banks—that would be prohibitively expensive and, for most banks, unnecessary. Even in 1988, the worst year of the bank crisis, only approximately 2 percent of U.S. banks failed. Therefore, regulators now help bridge the time between regularly scheduled examinations by combining off-site monitoring systems and additional examinations so that they have up-to-date evaluations of the financial condition of banks.

Off-site systems currently being used by bank regulators have several strengths. First, they are "current." That is, they are updated every quarter with new Call Report information. Second, they are far less intrusive than on-site examinations. This is very important.

---

[7] Andrew Albert, "Comptroller's Office Throws a Bash," *American Banker* 148 (November 4, 1983), 16.

[8] Barron H. Putnam, "Early-Warning Systems and Financial Analysis in Bank Monitoring," Federal Reserve Bank of Atlanta *Economic Review* 68 (November 1983): 6–12.

[9] Drew Dahl, Gerald A. Hanweck, and John O'Keefe, "The Influence of Auditors and Examinations on Accounting Discretion in the Banking Industry" (paper presented at the Academy of Financial Services conference, October 1995).

To achieve the same level of surveillance without these systems would require more on-site examinations and more staff. Third, these systems help regulators target examination resources efficiently. Institutions that show signs of financial distress can have their examination dates moved forward, or an institution can be contacted and asked to explain the changes observed. This also means that well-run and highly rated institutions will generally not be examined outside of the regular examination schedule. Fourth, today's off-site systems enable the failure models to be modified and updated with relatively few staff resources. Finally, whereas examinations focus on the current condition of the bank, off-site systems—which are current in terms of information—have the potential to identify high-risk characteristics that may increase the probability that a bank will fail.

Although the systems now in use function reasonably well, they have some weaknesses that generally stem from their complete dependence on Call Report data. For example, Call Reports do not note either the quality of management or management practices, as on-site examinations do, so the evaluation of management remains outside the realm of off-site systems. Likewise, under current methods, only on-site examinations look at individual loan files. A less-serious example of the problem with relying solely on Call Report data is that the accuracy of any of the models' data depends on on-site examinations (accordingly, the predictive power of the models decreases as the time between examinations increases). In addition, because of increased industry consolidation, only on-site examinations can determine the geographic loan concentrations of some banks.[10] Finally, because contemporary off-site models are used to assist in the examination process, they are "current condition oriented," which is their first strength, but for that very reason they do not measure the long-term risk in a bank—yet key aspects of changes in a bank's operations may take place as much as four or five years before a bank's crisis.

Discovering What a Forecasting System Can Do

To see why today's surveillance models work well in identifying a bank's current condition but not the risks a bank may face well into the future, researchers at the FDIC examined the characteristics of banks that failed and banks that survived over a five-year period. To examine how banks' condition changed over time, they constructed a data set consisting of all banks that existed in 1982 and *either* were still in existence in 1987 *or* had failed in 1986 or 1987 (banks that failed after 1987 or between 1983 and 1985 were excluded). The set of banks examined therefore contained two clear types: those that existed over the entire five-year period and never experienced failure, and those that existed at the beginning of the five-year period and failed during the fourth or fifth year.

[10] David Holland, Don Inscoe, Ross Waldrop, and William Kuta, "Interstate Banking—The Past, Present, and Future," *FDIC Banking Review* 9, no. 1 (1996): 1–17.

Four indicators of bank condition were examined: (A) equity ratio, (B) coverage ratio (equity plus reserves less delinquent loans, to total assets), (C) return on assets, and (D) nonperforming loans (see figure 13.1). In 1982, banks that would not fail during the next five years had an average equity ratio of 8.84 percent, while banks that would fail had a ratio 55 basis points lower (8.29 percent). This lower ratio is above the level that, under the risk-based system now in effect, is considered well capitalized. The coverage ratio, of course, was also lower for future failures: 6.57 percent versus 7.90 percent; so was the return on assets: 86 basis points versus 101 basis points. Nonperforming loans were slightly higher in the future failures: 2.3 percent of assets, versus 1.44 percent of assets in nonfailed banks. For all of the indicators, the average was worse for the future failures than for the survivors. However, these ratios would not in themselves be considered typical, or predictive, of banks that would fail, for the future failures also had good capital levels, decent earnings, and a low percentage of nonperforming loans.

With each passing year, the divergence between the healthy banks and the failed banks grew. By 1984, three years before failure, the equity ratios of the failing banks were 179 basis points lower than those of the nonfailed banks (6.85 percent versus 8.64 percent). The healthy banks had maintained a return on assets of 84 basis points, whereas the failures had fallen to -77 basis points. The future failures also showed the beginnings of large increases in their nonperforming loans, which had risen from 2.3 percent in 1982 to 5.05 percent in 1984.

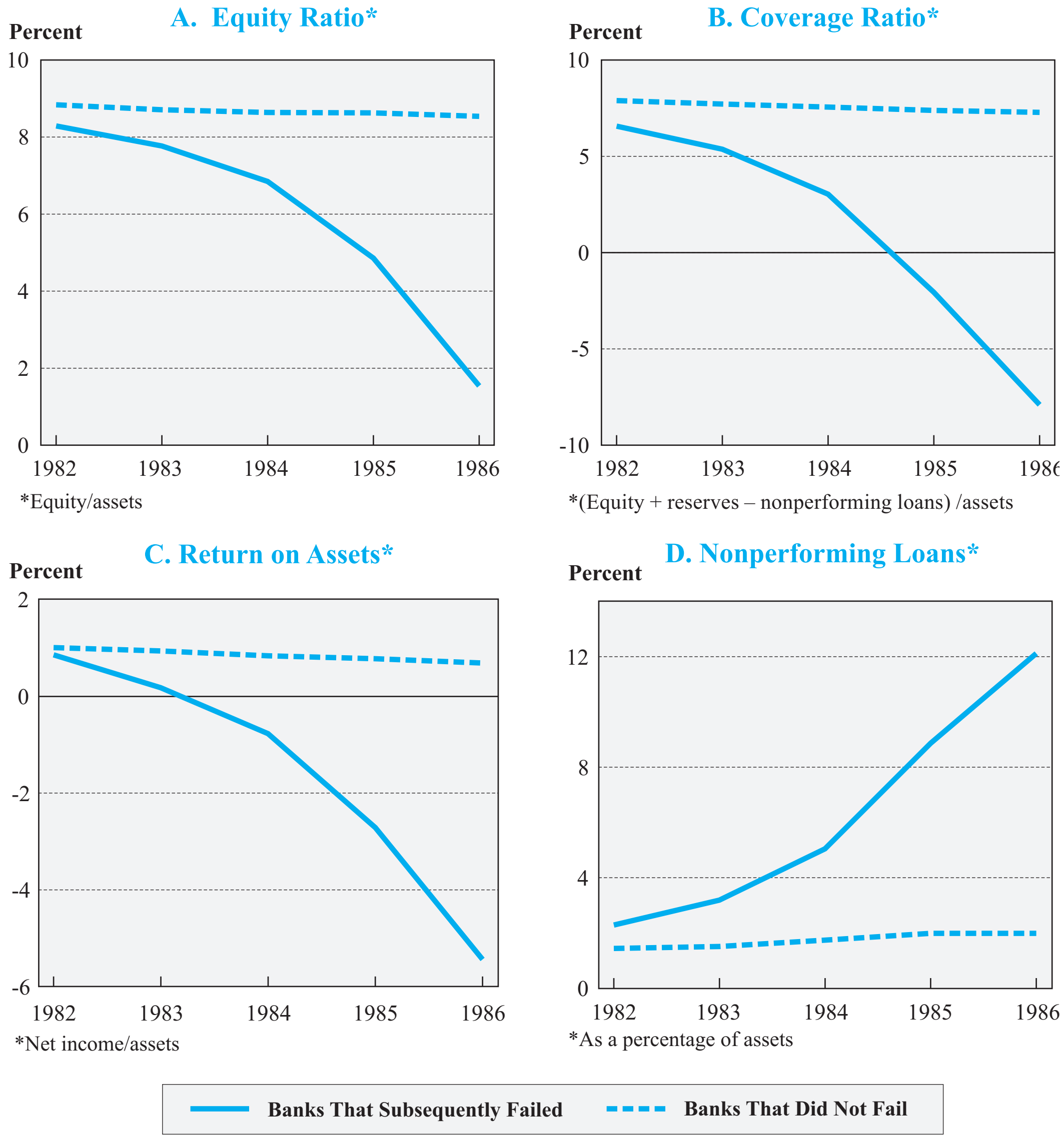
The data from 1985 demonstrate the wide differences that had developed between the two groups of banks. Equity at the healthy banks was virtually unchanged at 8.63 percent (compared with 8.64 percent in 1984), whereas at the future failed banks it had dropped 199 basis points to 4.86 percent. The failed banks' coverage ratio had fallen below zero (-2.06 percent); losses were accumulating rapidly, bringing the return on assets down to -2.71 percent; and the level of nonperforming loans had increased 76 percent to 8.87 percent of assets, above the average equity of three years earlier.

At the end of 1985, just before their failure, the failing banks are easy to identify. Their average equity was a very low 1.54 percent (healthy banks had 8.54 percent) and they were suffering enormous losses, with an average return on assets of -5.44 percent; nonperforming-loan ratios exceeded 12 percent. These data clearly show, therefore, that standard indicators of condition can identify banks that are already in financial distress but do not indicate which banks may *become* distressed.

Instead of looking at indicators of condition, if we look at the risk characteristics of the same banks over the same five-year period, we find a somewhat different pattern. Whereas the condition indicators for failed and surviving banks were very similar many years before failure, some of the risk indicators show wide differences several years prior to failure. The

Figure 13.1

Bank Condition Ratios for Failed and Nonfailed Banks, 1982–1986



Note: "Failed" means banks that existed in 1982 and failed in 1986 or 1987; "nonfailed" means banks that existed during the entire period and never failed.

four ratios used to measure risk in a bank were (A) the loans-to-assets ratio, (B) the asset growth rate, (C) the interest-and-fees-to-loans ratio, and (D) the salary-to-employee ratio (see figure 13.2).[11]

In 1982, in all four risk categories the surviving banks had lower average ratios than the failed banks. The surviving banks had a loans-to-assets ratio of 49.6 percent, a full 10 percent below the 59.8 percent ratio of the failed banks. Failed banks had an interest-income-and-fees-to-loans ratio that was almost 200 basis points above the ratio of the surviving banks (8.91 percent versus 6.97 percent). Failed banks were also growing slightly faster than the survivors: 13.9 percent per year versus 12.1 percent. And failed banks had salary-to-employee ratios that were 5.7 percent above those of surviving banks: $20,364 per employee for failed banks and $19,272 for survivors.

The pattern that developed over time for the risk indicators was very unlike the pattern for the condition indicators. For three out of four of the risk indicators, the difference between failed and surviving banks hardly changed at all. By the end of 1986 the failed banks had an average loans-to-assets ratio 12 percent higher than that of surviving banks (in 1982 the difference was 10 percent). The interest-and-fee-income ratio was still 200 basis points higher for failed banks than for survivors; and the failed banks' salary ratio—which in 1982 had been 5.7 percent higher than that of the surviving banks—was 4.8 percent higher ($24,637 for failed banks, $23,500 for survivors). The only ratio that demonstrated a dramatic difference over time was the asset growth rate. Over the entire period the asset growth rate for failed banks plummeted, going from a high of 13.9 percent in 1982 to 9.88 percent in 1984 and then to -5.5 percent in 1986, but the asset growth for surviving banks never fell below 8.8 percent.

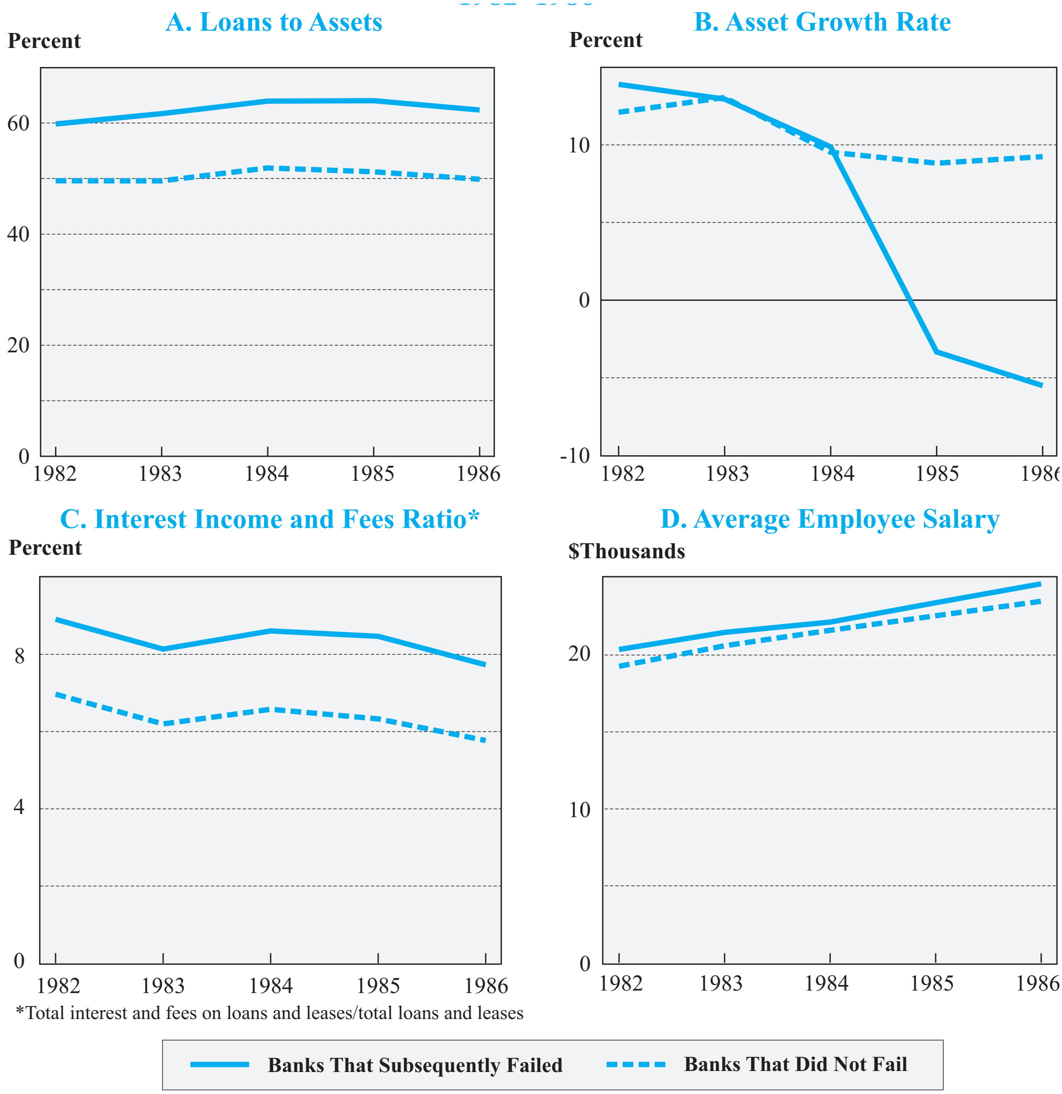
The condition indicators and risk indicators behave in such dissimilar ways (except for asset growth) that they are obviously measuring different aspects of banks. The current condition of a bank, as measured by the four condition indicators discussed above, can be viewed as the result of the risks the bank has accepted over a number of years. Exposure to excess risk can ultimately produce the conditions that cause failure. Exposure to risk involves the types of loans the bank issues or the type of business it chooses to enter, and in their day-to-day operations banks are continuously changing their risk exposure. Eventually such changes are reflected in the condition statements of the banks. If risk can be measured, it might be possible to see if banks that engaged in riskier practices failed at a higher rate than less-risky banks.

---

[11] The definitions of these risk ratios and explanations of what they measure are presented in table 13.1 (in the subsection entitled "Developing a Procedure").

Figure 13.2

Bank Risk Ratios for Failed and Nonfailed Banks, 1982–1986



Note: "Failed" means banks that existed in 1982 and failed in 1986 or 1987; "nonfailed" means banks that existed during the entire period and never failed.

Banks earn profits by accepting and managing risk. For example, when a bank issues a loan, the bank's management is making a conscious decision to accept the risk that the borrower will default. By issuing a large number of loans the bank can spread the risk of default over an entire portfolio. Borrower default is just one of the risks that bank management faces—and an important aspect of management's responsibilities is to establish the levels and types of risks the bank can accept, given management's ability to manage risk and the bank's ability to absorb the losses that may result. If the bank accepts too little risk, earnings will suffer, but if it accepts too much, it might face losses that would consume the institution's capital.

The types of risk a bank faces include credit risk, interest-rate risk, concentration risk, liquidity risk, and operating risk. Credit risk is the risk of default by a borrower. Interest-rate risk refers to the risk that an asset will lose value as interest rates rise or fall, or the risk that interest-rate changes will adversely affect income. Concentration risk refers to a situation in which a large percentage of assets are concentrated in one product or in one geographic area. This type of risk can flow from the very nature of the bank's business. For instance, small banks in agricultural communities are highly exposed to the risks of the agricultural economy. Likewise, specialized mortgage lenders are highly exposed to extreme changes in mortgage markets. Concentration risk can also occur when an institution undergoes rapid growth: the rapid growth results in the bank's having a high concentration of unseasoned loans, probably approved in a boom economy, or at least a benign one, but this high concentration of recent loans puts the institution at considerable risk when the economic environment worsens. Liquidity risk refers to potential difficulties in meeting cash demands from liability holders out of current assets. Operating risk is the risk of loss from mistakes and inefficiencies in the operation of the bank. A bank can fail from any one of these risks or from a combination of them.[12]

These risks may be magnified when bank management changes the institution's goals. For example, one particularly well-documented case is that of Continental Illinois (see Chapter 7). In 1976, acting on a report by the management consultants McKinsey & Co., the bank made very significant changes in its operating philosophy and decided to concentrate its lending in high-growth segments of the economy. In addition, to implement this strategy fully the bank "decentralized" its lending function and made loan approvals much easier to obtain.[13] In other words, the bank made a conscious decision to increase its risk profile. By concentrating lending in high-growth areas—that is, by lending into a "boom" sector—management increased the risk that loan defaults would result when the bust occurred. By reducing management controls for loan approvals, the bank also made it more

---

[12] George J. Vojta, *Bank Capital Adequacy* (1973).
[13] *Business Week* (October 21, 1982): 82.

likely that loans would go to financially weak firms. Not long after initiating these changes, Continental's senior management established a goal of growing to be one of the three largest commercial lenders in the nation.[14] Within two years after changing its goals, Continental had markedly increased its risk exposure.

Though it is difficult to detect differences in the financial condition of failing and surviving banks many years in advance of the failure, it may be possible to determine if failed and surviving banks have different risk characteristics. But even if it is possible to identify risk characteristics and therefore to identify a large percentage of eventual failures, it is nonetheless true that among banks with the same risk characteristics, a very high percentage may survive.

Thus, both accuracy and comprehensiveness are required if a system or model is to be judged effective. A failed-bank model might be calibrated so that a high percentage of its predicted bank failures actually fail, with a correspondingly low percentage of predicted bank failures that actually survive. This high accuracy, however, may not mean that the model identifies all, or even a majority, of the problem institutions. Alternatively, the model can "flag" a large percentage of the total number of banks as potential problems or failures, and although the probability that any individual bank will actually fail is low, a large percentage of failing institutions will nonetheless eventually be captured.

In statistics one quantifies these trade-offs by deciding what type of error one is willing to accept—Type I or Type II. A Type I error is an error one makes by rejecting a null hypothesis when the null hypothesis is in fact true, and a Type II error is an error one makes by accepting a null hypothesis when the alternative hypothesis is in fact true.[15] The trade-off between Type I and Type II errors is exemplified by the U.S. criminal justice system, in which a person is presumed innocent until proven guilty. In a criminal trial, the null hypothesis is that a defendant is not guilty. A Type I error occurs when an innocent person is found guilty (convicting the innocent). A Type II error occurs when a person who is guilty is incorrectly acquitted (acquitting the guilty). There is an obvious trade-off between the two types of errors. If one wants to have a very low Type I error (few innocents wrongly convicted), one usually accepts the fact that there will be a large Type II error (a large percentage of acquitted people will in fact be guilty). To minimize the occurrence of the Type I error, the courts require that there be evidence "beyond a reasonable doubt" in order to convict someone.[16] Likewise, if a small Type II error is desired (so that few people who are

[14] Ibid., 83.

[15] Richard W. Madsen and Melvin L. Moeschberger, *Statistical Concepts with Applications to Business and Economics* (1986), 360–65.

[16] In a civil case, the standard is the less-exacting "a preponderance of the evidence."

actually guilty are acquitted), then there is likely to be a very large Type I error (many innocent people will be judged guilty).

These trade-offs are inherent not only in statistical models but also in the bank examination system. All banks are examined within 18 months of the previous examination whether or not there is any evidence of a negative change in the bank's financial condition. The examinations are performed to capture the relatively few banks that have significant changes. Thus, in contrast to the criminal justice system, the bank examination process has a large Type I error: many healthy banks are examined so the regulators can find the few that had negative changes. These trade-offs are important to keep in mind when one considers the various surveillance systems.

Real-World Obstacles to Forecasting

For several reasons, it is difficult to identify future problem banks even when the effort is made to identify risk factors. The life cycle of problem banks is such that in its early years, future problem banks cannot yet be clearly distinguished from other banks. In addition, both the economic environment and the financial process are dynamic and not easily modeled by the forecasting tools available.

The Life Cycle of a Bank Failure

In interviews with bank and thrift regulators, rapid loan growth was identified again and again as a precursor to failure. Whether or not loan growth is the primary risk in which banks engage, one regulator's description of a three-phase process by which rapid loan growth evolves into a major problem does a good job of laying out the long-term nature of the development of a bank's financial distress.

In the first stage, there is rapid loan growth; loan concentrations emerge, and lending is aggressive (internal controls in the growth areas are weak, and underwriting standards are lenient). The increased lending may be, but is not always, funded by a volatile lending source. This growth could occur throughout the entire institution or within a specific asset type. If the growth is in a specific asset type, the increase could stem either from growth in concentration in a loan category or from a shift into a new activity, with subsequent growth. If the rapid growth draws the attention of the relevant regulator, management usually points to the excellent earnings and contribution to capital that the growth has provided. This stage of the development of the problem can take up to two years.

In the second stage, the institution has rising loan-quality problems. Associated expenses may far exceed industry averages. Nonrecurrent sources of income are used to maintain the same level of profits that existed during the growth phase. Eventually profits begin to decline, and inadequate reserve levels become apparent. At this point the bank may be "loaned up" (that is, have a high loans-to-assets ratio). Management may still believe that the problem is manageable. This stage may take an additional one to two years.

In the final stage, deteriorating asset quality is a serious problem. The institution is incurring large loan losses, and charge-offs have increased. If the institution is large, the capital markets have recognized that the institution has inadequate loan-loss reserves and are unwilling to provide fresh capital. At this point, major changes in the bank's operations are necessary. Dividends may be cut, expenses (mostly personnel) are slashed, and assets are sold to cover charge-offs and operating expenses (especially in larger institutions). This crisis phase may last up to a year and results either in the failure of the bank or, if dramatic and fundamental changes are made, in its eventual recovery.

As this account of the life cycle of failure makes clear, only in the course of years do changed behavior and the acceptance of greater risk lead to financial distress or failure. After all, neither growth itself nor most other risk taking is necessarily bad for a financial institution. Banks earn their income by assuming risk; to increase risk through growth can therefore be a sound strategy. Such a strategy would ideally be accompanied by increases in capital as a buffer against higher losses, maintenance of high underwriting standards, and attention to proper risk management—in other words, by prudent management of the institution's growth. Moreover, regardless of whether the increased lending is prudent, ill timed, or very risky, the growth will generate added revenue from increased loan fees and interest income. In addition, because these are all new loans, initially there are no delinquencies and no loss charge-offs, so that the growth is almost always accompanied by growth in income and capital (assuming retained earnings). Only over time do the effects of growth or other risk taking—whether these effects are good or bad—become apparent. This long lead time before problems appear makes it difficult to identify future problem banks accurately.

The Dynamics of the Economic Environment

Long lead times are not the only problem encountered in forecasting failures. There are two others.

One is that economic conditions, both regional and national, change over time, but the changing nature of economic conditions is not built into failure forecasts. All failure forecasts are based on financial profiles of banks, indicating whether a bank has the characteristics of other banks that have failed. This seems relatively straightforward. If it is found that failed banks have low capital levels, high percentages of nonperforming assets, and poor earnings, then nonfailed banks with similar financial profiles should be considered probable failures. Embedded in this type of analysis, however, is the underlying assumption that the set of economic conditions under which the failures occurred will not change. Without explicit economic variables in a model, the forecasts for future failures assume the same economic environment as the one in which the actual failures occurred: the then-current interest-rate environment, the particular real estate market, and the same general nationwide economic health. But if economic conditions change, as they always do (for example, there may be a recession or a dramatic interest-rate change), the number of actual failures (or CAMEL rating downgrades) can substantially diverge from the forecasts.

The Dynamics of the Financial Process

Finally, forecasting is difficult because normal economic models assume linearity, but as the three-stage life cycle shows, the financial process that leads to failure is inherently nonlinear. Failure is a rare event, and only extreme behavior eventually causes a bank to fail. For an analogy, consider the situation of people who are overweight (assuming that excess weight is bad for a person's health): if overweight people continue to gain weight their health will worsen, and if they lose weight their health will improve—but if they lose *too much* weight, their health will again suffer. Many aspects of bank risk taking can be thought of in the same way: too much growth can result in financial distress, but too little may threaten the bank's long-term financial viability. This "too much or too little" phenomenon makes the financial process nonlinear; hence, both very high growth and very low growth may be "risky." For that reason, economic models that attempt to capture the specific dynamics of the financial process are unstable and lumpy, and do not isolate the risks of failure.

Analysis by Risk Groups

To isolate these risks, contingency table analysis is needed in which the specific dynamics of the process are ignored and one looks at "levels" of risk or risk groups to classify banks or people (the underlying dynamics of the process, nevertheless, are always present). Analysis by risk groups is most common in epidemiological studies. For example, a person who smokes has twice the risk of having a heart attack compared with a person who does not smoke. The risk of a heart attack is also double for a person who has high blood pressure or high blood-cholesterol levels. In addition, these risk factors are multiplicative: if a person has two factors, the risk of a heart attack increases four times; if all three factors are present, the risk increases eightfold.[17] For banks it may be possible to determine risk factors in a similar manner—in other words, to develop nonlinear models. The two subsections that follow give details of an attempt to do that.

---

[17] NIH Pub. No. 93-2724, rev. October 1992, National Heart, Lung and Blood Institute, National Institutes of Health.

Developing a Procedure

In connection with heart attacks, the levels for "high blood pressure" or "high cholesterol" have already been determined. In contrast, for banks the levels for risk factors have not yet been identified. We assume, however, that risk increases when the risk measure increases. The goal in analyzing risk measures is to find the set of variables that has the greatest predictive power for determining which banks will fail.

A group of researchers at the FDIC chose nine measures of risk to study and eventually used eight of them (see table 13.1). To determine how these measures of risk predict failure individually and as a set, the researchers divided each measure into five risk groups (quintiles) from high to low, using the data for the years 1980, 1982, 1984, 1986, and 1988. For each year studied, banks that never failed were separated from banks that failed four or five years later (all other banks that existed for only part of the five-year period were excluded from the study, as is explained in more detail below). Both groups of banks in each period were then analyzed to determine which risk measures were the best long-range predictors of failure (the details of the analysis also appear below).

A brief summary of the results of the analysis appears here (a fuller presentation appears in the next subsection). Among this group of variables, the best long-range predictor

Table 13.1

Ratio Measures of Bank Performance

Identification of Variable	What the Variable Measures
Loans-to-assets ratio	Liquidity and risk. The higher the ratio, the greater the amount of the bank's total portfolio that is subject to default risk.
Deposits over $100,000 (large deposits) to total liabilities*	The use of larger deposits to fund assets. These deposits may be more volatile than fully insured deposits.
Return on assets	The bank's profitability. Low ROA may encourage risk taking by the bank. High ROA may indicate high-risk lending to increase profits.
Asset growth from previous year	Risk of growth.
Loan growth from previous year	Risk of growth.
Operating expenses to total expenses	Management's control of expenses. Higher expenses are assumed to be an indicator of loose controls.
Salary expenses per employee	Management's control of expenses.
Interest on loans and leases to total loans and leases (interest yield)	The average income of loans. High yields might indicate that the bank is originating high-risk loans.
Interest and fee income to total loans and leases (interest and fees to loans)	Income. The addition of fees to the variables may catch firms that are loading up on fee income.

* This variable was eventually dropped (see the discussion below about banks in Texas).

of failure is a bank's loans-to-assets ratio. This result appears to be consistent across all years and all regions. In all five years studied, approximately 50 percent or more of the failures come from the top loans-to-assets quintile. In the last three periods (1984 through 1988), if banks in that quintile are excluded, then the banks in the highest return on assets (ROA) risk group are the best predictor of failure.

The evidence is strong that the basic pattern of bank distress and failure as set forth by the regulators and presented above is valid. Banks that eventually become troubled do undertake risky business strategies several years before their financial condition deteriorates. But even if it turns out to be possible to identify these risky strategies, it may still be very difficult to identify which banks within a risk group will fail and which will survive. In addition, the predictions have a large Type II error: although the procedure identifies the quintile that contains a very large percentage of the failures, more than 95 percent of all the banks in the quintile never fail.

Contingency Table Analysis: Methodology and Results

The data for the study were constructed from all BIF-insured institutions (banks and savings banks) that existed in the beginning year and *either* did not ever fail (then or later) *or* failed four or five years from the beginning date. Thus, the study excludes banks that existed in the beginning year and (a) failed before the fourth year, (b) were merged out of existence during the period, or (c) failed subsequently; and it also excludes all de novo banks created during the period. The reasons for the exclusions were that banks that failed or merged in the interim period were not in the sample long enough to be studied, nor were de novo banks, and banks that failed subsequent to the period under study were excluded to ensure that each sample had clearly defined groups of survivors and failures.

So that an epidemiological approach could be used, a contingency table analysis was performed on each year's data. First, a logit regression was performed on each variable, where the dependent variable was whether the bank failed or did not fail (1 or 0). The variable with the highest predictive power for failure was determined by a Chi-Square test score for each regression. The coefficients for each quintile grouping of the variable were then compared, and a Chi-Square test was performed to determine which quintile or group of quintiles was the best predictor of failure. The split of the quintiles created a "high-risk" group and a "low-risk" group. The analysis was then repeated on both of the two groups to determine the next-best predictor of failure in each group. This procedure was repeated for each subgroup until the cells became too sparse (the number of failures was too low) to analyze (see figure 13.3).[18]

---

[18] The procedure used was not complicated but was very time-consuming. Thus, it was important to keep the number of analyzed variables at a reasonable level.

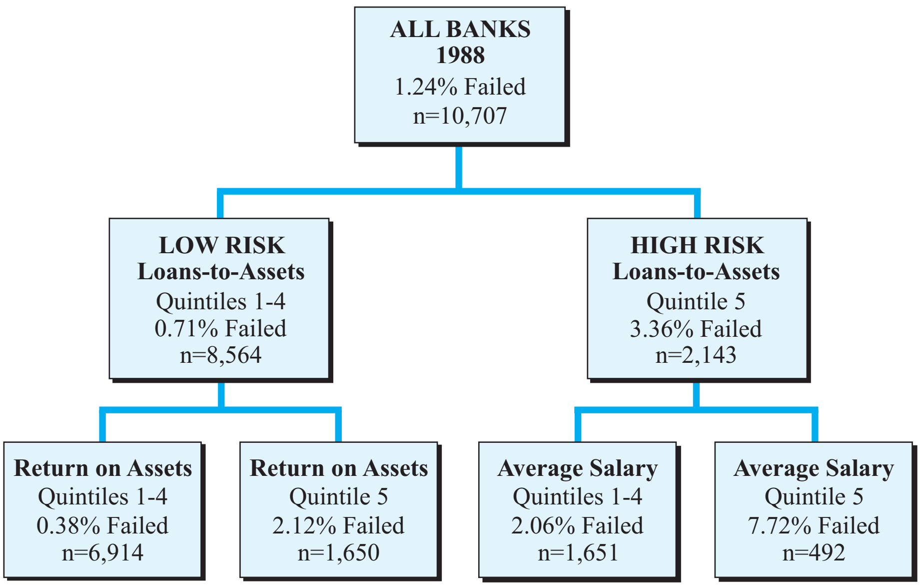
The five study periods began in 1980 and spanned ten years of failures, from 1984 to 1993. Included were 1,193 failures. Not included were 300 failures that occurred during the period but were excluded from the study because they fell into one of the following groups: (1) banks that did not exist for at least four years, (2) banks that were taken over under a "cross-guarantee" subsequent to the passage of the Financial Institutions Reform, Recovery, and Enforcement Act of 1989, (3) banks that were closed primarily because of fraud, and (4) subsidiary banks of First Republic and First City that had composite CAMEL ratings of 1 or 2 as of closing (similar to cross-guarantees).

An examination of the relationship between the nine variables and the failures over the five different periods reveals that banks in the highest loans-to-assets quintile had the highest probability of failure for the periods beginning in 1980, 1982, and 1988 and the second-best "high-risk" probability in 1984 and 1986. In those two years the best predictor was the large-deposit ratio.

However, because the large-deposit ratio did not show up as either a primary or a secondary indicator in 1980 or 1982, there was concern that it might not be an indicator of

Figure 13.3

Procedure Used in Contingency Table Analysis



volatile funding. This issue arose apropos of banks in Texas—the predominant state in the Southwest—which was a unit-bank state. The lack of branching might have forced Texas banks to rely more heavily on large deposits than did banks in states with branching. In periods of high growth, the inability to produce deposits through a branch system might also produce high ratios of large CDs. It was hypothesized that the large-deposit ratio might be a function of the particular region rather than an actual risk measure. To test the theory, the researchers examined the distribution of Texas banks' large-deposit ratio and found that Texas banks were extraordinarily concentrated in the high quintiles of large deposits. In 1980, 63 percent of Texas banks were in the two highest quintiles. During the next four years, assets in Texas banks grew 66 percent (from $118 billion to $198 billion),[19] and by year-end 1984, 84 percent of Texas banks were in the two highest quintiles of the large-deposit ratio (58 percent were in the highest quintile). In 1986 the comparable figures were 89 percent and 68 percent. It appears that Texas started the 1980s with a higher-than-average number of banks with a high percentage of large deposits, and banks in that state disproportionately used large deposits to fund asset growth. Thus, large deposits indicated a high probability of being a Texas bank rather than being an indicator of risk, so large deposits were dropped from the list of variables.

Once large deposits were excluded, the loans-to-assets ratio was always the best predictor of future failure. Being in the highest loans-to-assets quintile more than doubled a bank's probability of failure (see table 13.2). More important, after 1980 more than 50 per-

Table 13.2

Probability of Failure When a Bank Appears in the Highest-Risk Category

	Aggregate Failures		**Highest Loans-to-Assets Quintile**		
Beginning Year	Probability of Failure (Percent)	Number of Failures	Probability of Failure (Percent)	Number of Failures	**Increased Probability of Failure from Total Population to Banks in Highest Loans-to-Assets Quintile** (Percent)
1980	1.51	184	3.62	88	140
1982	2.45	291	6.75	160	175
1984	2.89	332	8.20	188	184
1986	2.25	253	6.46	145	187
1988	1.24	133	3.36	72	171

[19] FDIC, *Statistics on Banking: A Statistical History of the United States Banking Industry, 1934–1994,* vol. 2 (1995), E–545.

cent of the total failures for each cohort of banks came from the highest loans-to-assets quintile.

As noted, in the contingency table analysis the banks were split into two groups, the "high-risk" group (in this case, the banks in the highest loans-to-assets quintile) and the "low-risk" group (all other banks), and the calculations described above were repeated so that the next-greatest indicators of risk could be found.[20] This second-level analysis for the high-risk group did not yield a consistent pattern for second-level predictors. In 1980 and 1982, the interest-and-fee-income ratio was the best second-level predictor; in 1984, 1986, and 1988 the second-level predictors were, respectively, asset growth, return on assets, and average salaries. This result is discouraging, for it indicates that the relationship between the second-level risk indicators and failure is unstable (see table 13.3).

If a bank was not in the highest-risk quintile, that did not mean the bank had no risk of failure. A little under half of all banks that failed were not in the high loans-to-assets quintile, so it may be useful to see if the remaining banks that failed had any identifiable risk characteristics. For banks that were not in the high-risk loans-to-assets quintile, the best predictors of failures were loan growth in 1980, interest yield in 1982, and ROA in 1984, 1986, and 1988. The risk indicators for the so-called low-risk groups (that is, all groups except the highest-risk quintile) performed quite well. They identified a very large percentage of the remaining failures, particularly in 1986 and 1988, when being in the highest ROA quintile identified 57 percent of the remaining failures. If the high-risk and low-risk groups

Table 13.3

Probability of Failure When a Bank Appears in the Highest- and Second-Highest Risk Categories

	Highest Loans-to-Assets Quintile			Subset of Loans-to-Assets Quintile	
Beginning Year	Probability of Failure (Percent)	Percent of Total Failures	**Second-Level High-Risk Ratio**	Probability of Failure (Percent)	Percent of Total Failures
1980	3.6	47.8	Interest and loan fees	7.2	31.5
1982	6.8	55.0	Interest and loan fees	11.6	35.4
1984	8.2	56.6	Asset growth	12.6	28.0
1986	6.5	57.3	Return on assets	12.0	27.3
1988	3.4	54.1	Average salary	7.7	26.3

[20] After the first "high-risk" group was identified, the remaining banks were not redistributed into new, equal quintiles. Rather, they were left in the original quintile distribution, with the already identified "high-risk" banks removed.

were taken together in 1988, they would contain more than 80 percent of the study group failures (see table 13.4).

From these results, one can infer that it may be possible to identify groups or populations of banks with a high probability of containing a high proportion of future failures, or that it may be possible to identify large populations of banks with a very low probability of failing in the future. Thus, the analysis described above has limitations that must be considered. First, to identify 80 percent of the failures, the contingency analysis "flagged" 35 percent of the entire study population for 1988: 2,143 banks in the loans-to-assets quintile and 1,650 banks in the ROA quintile, or a total of 3,793. The entire study population for that year consisted of 10,707 banks, 133 of which failed, and 107 of the failures (80 percent of 133) were in two identified risk groups. The two identified risk groups also contained 3,686 banks that did not fail, or approximately 97 percent. In addition, in 1988 there was no way to identify *which* 3 percent would fail in 1992 or 1993. Nor would identification have been much easier if only the highest-risk loans-to-assets banks had been identified. In the 1988 cohort approximately 96 percent of the high-risk loans-to-assets quintile survived, and in the 1984 cohort (the one with the highest number of failures), 92 percent survived. Second, to differentiate clearly between failures and survivors, the analysis was performed on a subset of all banks, but the exclusion of some banks from the analysis might have introduced measurement errors. Third, the lack of consistency in the secondary risk factors may mean that the industry changes so rapidly that supervisory attention could be diverted to monitoring diminishing risks instead of identifying emerging risks.

Table 13.4

Probability of Failure in "Low-Risk" Banks

(Banks Not in the Highest-Risk Group)

	Highest Loans-to-Assets Quintile			**"Low-Risk" Failure Indicator**	
Year	Probability of Failure (Percent)	Percent of Total Failures	**High-Risk Indicator for "Low-Risk" Group**	Probability of Faillure* (Percent)	Percent of Remaining Failures †
1980	3.62	47.8	Loan growth	2.32	41.7
1982	6.75	55.0	Interest yield	3.76	40.4
1984	8.20	56.6	Return on assets	3.96	45.1
1986	6.46	57.3	Return on assets	3.74	57.4
1988	3.36	54.1	Return on assets	2.12	57.4

* This is the probability of failure in the remaining 80 percent of banks that are not in the high-risk loans-to-assets quintile.

† Excludes failures in the high-risk loans-to-assets quintile.

The FDIC's Growth-Monitoring System (GMS)

The contingency analysis—attempting to identify the interactions within a set of risk groups in order to find a way to predict future failures—feeds into and seeks to improve the FDIC's growth-monitoring system (GMS). GMS was developed during the mid-1980s and was designed to detect the initial stage in the life cycle of failing banks—the rapid-growth stage. The system's premise is that rapid growth in total assets (or loans) represents a risky activity of which bank supervisors should be aware. Growth-related risk can come in at least two areas, loans and bank management: there may be increased loan concentrations in risky areas, and there may be management lapses such as lowered underwriting standards, increased reliance upon volatile funding, or a general weakening of internal controls in order to facilitate rapid growth. Banks that GMS identifies as rapid-growth institutions in these two areas are flagged for off-site review and may receive increased supervisory attention.

The system is based upon the levels and quarterly trends of five summary measures. These include two growth rates (for total assets and for loans and leases) along with three financial ratios (as percentages of assets): loans and leases, plus securities with maturities of five years or more; volatile liabilities; and equity capital.[21] The system measures both the levels and the trends (growth) of the three financial ratios in addition to asset growth and loan growth, for a total of eight terms. Banks' percentile rankings are computed quarterly for each of the eight terms; all percentile rankings are relative to a bank's Uniform Bank Performance Report (UBPR) peer group (see appendix B). There are 25 UBPR peer groups based on asset size, location in a metropolitan area or a nonmetro area, and number of branch offices. These eight percentile rankings are subsequently weighted in a two-step process, and the weighted percentile rankings of the eight terms are then summed to give a GMS score (see table 13.5).

Composite GMS scores are evaluated separately for two groups of banks. The first group is composed of banks whose quarterly asset and loan growth rates were 5 percent or more (high-growth banks). For all high-growth banks, composite GMS score percentile rankings are computed. Banks in the highest composite GMS score percentiles—currently the 95th to 99th percentiles—are "flagged" for off-site review. Supervisors may also review banks beneath the 95th percentile, particularly those with poor CAMEL ratings. The second group is composed of banks with quarterly asset and/or loan growth under 5 percent (low-growth banks). These low-growth banks' GMS scores and related information are available for review by regional office examiners in the GMS system.

---

[21] Volatile liabilities are defined as the sum of the following: time deposits of $100,000 or more, deposits in foreign offices, federal funds purchased and repurchase agreements, demand notes issued to the U.S. Treasury, and other liabilities for borrowed money.

Table 13.5
Hypothetical GMS Score Computation Example

	Trend Weight		Trend Percentile		Ratio Weight		Ratio Percentile		Raw Score		Weight		Score
Asset growth	0.60	x	98	+	-	x	-	=	58.9	x	0.67	=	39.4
Loan growth	0.60	x	99	+	-	x	-	=	59.4	x	0.00	=	0.0
Loans and securities/assets	0.60	x	98	+	0.40	x	82	=	91.6	x	0.11	=	10.1
Volatile liabilities/assets	0.60	x	96	+	0.40	x	86	=	92.0	x	0.11	=	10.1
Equity/assets	0.60	x	98	+	0.40	x	85	=	92.8	x	0.11	=	10.2
						Composite GMS Score							**69.8**

The next subsection contains a detailed review of GMS's predictive abilities as many as five years before deterioration in banks' financial condition. The findings can be summed up as follows: GMS appears to perform the function for which it was designed. The system identifies a group of banks that have a higher-than-average risk of failure, and may do so up to four years before failure. When a standard failure estimation technique is used, the GMS score has also been found to be a significant long-term predictor of failure in three out of four time periods. In addition to predicting failure moderately well, GMS has been a better-than-average predictor of CAMEL downgrades two to three years in advance of the event.

No significant changes have yet been made to the system. However, marked and significant improvements have been suggested for each stage of the process (these suggested changes are also detailed in a later subsection). Major proposed improvements include a new weighting scheme for the GMS score, new variables for inclusion in the score, better methods of constructing growth variables, and use of peer groups not based on the UBPR groupings. If all of these suggested changes were made, they could increase the percentage of banks accurately identified as future problems and could decrease the percentage of banks incorrectly identified as future problems (in other words, the changes should decrease Type II errors).

Effectiveness

If GMS is effective, risk detection should occur well before there are adverse changes in banks' financial condition. Therefore, researchers evaluated GMS's predictive abilities as much as five years before deterioration by comparing (1) GMS composite scores with future bank-failure rates, and (2) GMS score percentile rankings with changes in banks' composite safety-and-soundness (CAMEL) ratings.

Early detection of failure. To review the first relationship (between GMS composite scores and future bank-failure rates), the researchers began by computing GMS composite scores for the last calendar quarter of each year between 1984 and 1994. Next they ranked banks' GMS scores into deciles, and compared bank failures occurring in subsequent years (between one and five years after scores were assigned) across GMS score deciles. For computational simplicity, open-bank assistance transactions were excluded from the group of failed banks. The analysis showed that banks in the lowest GMS score decile usually failed at the highest rates during the two years immediately after scores were measured. The situation was reversed, however, for failures occurring between three and five years after scores were computed: the long-term failure rates were generally higher for banks in the highest GMS decile. These results are in agreement with the life-cycle profile for failing banks.

For example, among banks ranked by GMS score deciles for December 1984, approximately 39 percent of the banks that failed during the next year were in the lowest (first) GMS score decile (see table 13.6). The proportion of 1986 failures in the lowest GMS decile was also high, at approximately 19 percent.[22] Failures in subsequent years were more frequent for the highest (tenth) GMS score decile: banks in the tenth GMS decile accounted for approximately 21 percent of 1987 failures, 20 percent of 1988 failures, 22 percent of 1989 failures, and 20 percent of all failures occurring between 1990 and 1995. Similar results (not presented here) were obtained for other GMS score ranking years.

Exam rating changes. The reason for targeting banks for inclusion on an off-site review list is that they may be undergoing rapid changes in condition. Thus, the second test of the usefulness of GMS was the relationship between GMS score percentile rankings and subsequent changes in banks' composite safety-and-soundness (CAMEL) ratings—more specifically, changes in CAMEL ratings measured between banks' most recent CAMEL rating as of the date of the GMS ranking and the examination subsequent to the ranking date (the analysis used examination ratings over a period of two years before and two years after the date of the GMS score ranking). Those tests (not presented here) revealed no consistent relationship between banks' GMS rankings and changes in CAMEL ratings, either downgrades or upgrades. That result was not unexpected, however, given the previous results for failure rates. The evidence from failure rates and GMS rankings indicates that rapid growth is related to failure three or more years afterward. To the extent that CAMEL ratings are concurrent as opposed to leading indicators of condition, deterioration in CAMEL ratings would not be expected in the near term. Rather, deterioration in CAMEL ratings among banks in the highest GMS decile is expected three or more years after rapid growth occurs.

[22] These results are consistent with those reported in Chapter 12.

Table 13.6

Bank Failures by GMS Score Ranking and Failure Year

(Number and Percent of Year's Failures)

December 1984 GMS Decile	Never Failed		Failed 1 year Later		Failed 2 Years Later		Failed 3 Years Later		Failed 4 Years Later		Failed 5 Years Later		Failed More Than 5 Years Later		Total	
1	1,318	9.74%	42	38.89%	26	18.98%	24	13.48%	16	8.79%	12	6.52%	28	8.00%	1,466	9.99%
2	1,363	10.07	15	13.89	22	16.06	19	10.67	10	5.49	13	7.07	25	7.14	1,467	10.00
3	1,364	10.08	6	5.56	18	13.14	16	8.99	15	8.24	22	11.96	26	7.43	1,467	10.00
4	1,376	10.17	7	6.48	12	8.76	12	6.74	14	7.69	18	9.78	28	8.00	1,467	10.00
5	1,389	10.27	5	4.63	6	4.38	13	7.30	16	8.79	10	5.43	28	8.00	1,467	10.00
6	1,373	0.15	4	3.70	10	7.30	18	10.11	9	4.95	17	9.24	36	10.29	1,467	10.00
7	1,367	10.10	6	5.56	14	10.22	10	5.62	15	8.24	21	11.41	34	9.71	1,467	10.00
8	1,364	10.08	9	8.33	9	6.57	15	8.43	26	14.29	14	7.61	30	8.57	1,467	10.00
9	1,347	9.96	5	4.63	14	10.22	14	7.87	24	13.19	17	9.24	46	13.14	1,467	10.00
10	1,269	9.38	9	8.33	6	4.38	37	20.79	37	20.33	40	21.74	69	19.71	1,467	10.00
Total	**13,530**	**100.00%**	**108**	**100.00%**	**137**	**100.00%**	**178**	**100.00%**	**182**	**100.00%**	**184**	**100.00%**	**350**	**100.00%**	**14,669**	**100.00%**

To test the latter hypothesis, the researchers looked at the relationships between GMS rankings at a given year-end and changes in CAMEL ratings assigned one and three years later—for example, the relationships between GMS rankings at year-end 1984 and changes in CAMEL ratings for exams given in 1985 and 1987 (see table 13.7). Nearly 16 percent of CAMEL downgrades occurred among the highest GMS decile group, a higher proportion than for any other decile. Moreover, for most examination years considered between 1984 and 1995, the proportion of downgrades generally rose with decile rankings. These results support the previous relationships between GMS rankings and future failure rates; they also support GMS's use as a leading indicator of bank risk.

Statistical significance of the results. What the previous analysis does not test is the statistical significance of the relationships between GMS score rankings and subsequent changes in banks' condition. To test the statistical significance of GMS scores in measuring bank risk, the researchers included GMS scores in standard models of bank-failure prediction.

Logit Model Methodology

Logit estimations of the relationships between banks' financial condition and the incidence of failure were obtained with the use of year-end financial data, actual failures, and assistance transactions during the subsequent two years (see appendix A). Equation 1 is the basic model:

Table 13.7

Comparisons of Exam Ratings as Assigned in 1985 and 1987

(Number and Column Percent)

December 1984 GMS Decile	CAMEL Upgraded Next Exam		CAMEL Downgraded Next Exam		No CAMEL Change Next Exam		Missing Because of Mergers, etc.		Total	
1	248	19.89%	93	8.73%	323	9.70%	802	8.88%	1,466	9.99%
2	159	12.75	81	7.61	357	10.72	870	9.64	1,467	10.00
3	130	10.43	100	9.39	351	10.54	886	9.81	1,467	10.00
4	121	9.70	81	7.61	361	10.84	904	10.01	1,467	10.00
5	117	9.38	96	9.01	333	10.00	921	10.20	1,467	10.00
6	114	9.14	106	9.95	361	10.84	886	9.81	1,467	10.00
7	97	7.78	98	9.20	333	10.00	939	10.40	1,467	10.00
8	95	7.62	111	10.42	295	8.86	966	10.70	1,467	10.00
9	96	7.70	129	12.11	296	8.89	946	10.48	1,467	10.00
10	70	5.61	170	15.96	319	9.58	908	10.06	1,467	10.00
Total	**1,247**	**100.00%**	**1,065**	**100.00%**	**3,329**	**100.00%**	**9,028**	**100.00%**	**14,669**	**100.00%**

Equation 1

Likelihood of failure(i, next two years) = c0
+ c1(Capital, loss reserves)(i)
+ c2(Loans past due 30–89 days)(i)
+ c3(Loans past due 90 days or more, non-accrual loans, repossessed real state)(i)
+ c4(3-year mean operating income)(i)
+ c5(3-year standard deviation in operating income)(i)
+ c6(Examination interval, normalized)(i)
+ c7(Most recent capital rating)(i)
+ c8(Most recent asset rating)(i)
+ c9(Most recent management rating)(i)
+ c10(Most recent earnings rating)(i)
+ c11(Most recent liquidity rating)(i)
+ c12(Average salary/employee)(i)
+ c13(Multibank holding co. dummy)(i)
+ c14(Log of bank assets)(i)
+ c15(GMS score current year-end)(i)
+ c16(GMS score prior year-end)(i)
+ c17(GMS score 2 years prior)(i)
+ c18(GMS score 3 years prior)(i)
+ c19(GMS score 4 years prior)(i)
+ e(i,t)

Models in the form of equation 1 had previously been tested and were found to be fairly accurate failure-prediction models.[23] Results from this model show that the lower a bank's GMS score in the most recent period, the higher the probability of failure. That result is consistent with those found earlier. Moreover, the relationship between the most recent GMS score and failure was statistically significant for four of the five estimations. The results for lagged GMS scores were not as consistent, however. The expectation had been that high lagged GMS scores would be positively related to failures, as above, but in fact lagged GMS scores were sometimes—but not consistently—significantly and positively related to the likelihood of failure.

[23] Gerald A. Hanweck, Gary Fissel, and John O'Keefe, "A Comparative Analysis of Modeling Methodologies of Financially Distressed Banking Firms" (paper presented at the Financial Management Association Meetings, October 1995).

Proposals for Improvement

Several proposals have been made to refine GMS. These are described below.

Distinguishing types of growth. GMS does not distinguish between two important types of bank growth—increases in assets through existing offices (internal growth) and growth in assets through mergers, acquisitions, and consolidations (external growth). It may be that the risk profiles of banks are different for the two types of growth. The researchers hypothesized that external growth is less risky than internal growth. Internal growth may require more "new business" for the bank in terms of customers, markets, and products/loan types, whereas mergers and consolidations are more likely to involve the acquisition of seasoned loans from target banks. Moreover, because of regulatory limits on geographic expansion of banks, many mergers occur between former competitors in a single market. Rapid growth in "new business" or unseasoned loans may be considered risky for several reasons. Information about new customers and new markets may be limited, and underwriting standards may be loosened as a way to expand business in existing markets.

To test for the importance of these factors in assessing the riskiness of growth, the researchers used a simplified version of the bank-failure prediction model (equation 1) and included a control variable for merger-related growth. The variable of interest is a dummy variable set equal to 1 if the bank was involved in a merger, acquisition, or consolidation during the quarter its GMS growth score was measured, and zero otherwise. Logit estimations show that growth by mergers was negatively related to the likelihood of failure; however, the coefficients for the merger dummy were usually not statistically significant. This negative relationship is consistent with the hypothesis about internal versus external growth.

Modifying the peer groups. Banks' financial performance often differs systematically across industry segments, so some form of peer ranking is needed in GMS. GMS puts banks into peer groups based upon the UBPR standards. As mentioned above, the 25 UBPR (Uniform Bank Performance Report) peer groups distinguish banks on the basis of asset size, location in metropolitan area, and number of branch offices.[24]

Some form of asset size grouping would appear to be necessary. For example, small banks that do not have easy access to direct financial markets for equity often rely on retained earnings for equity funding. This lack of flexibility in equity finance is a reason that small banks hold large amounts of excess or buffer capital relative to regulatory capital requirements. Large banks, however, do not suffer from financial diseconomies and may rely on new equity issues for additional capital. For those and other reasons, banks' capital-to-assets ratios generally decline as asset size increases. However, it is less clear whether lo-

[24] See UBPR peer definitions in appendix B, Table 13-A.6.

cation in a metro or nonmetro area and number of branch offices have significance in assessing risk. With the advent of interstate branch banking in 1997, these latter two criteria would seem to have become particularly irrelevant.

Several alternative peer-group designations were tested, and the results of risk detection based upon revised GMS score rankings were compared with those based on the original 25 UBPR groups. One of the more promising alternatives tested was peers formed on the basis of seven geographic regions and two asset-size ranges (assets over and under $1 billion). Preliminary results (not presented here) indicate that rankings of GMS scores using this peer grouping performed marginally better than the 25 UBPR peer groupings in detecting banks likely to have CAMEL rating downgrades. Although the relevant peer groups depend in large measure upon the ratios used in scoring, it seems likely that a simplified peer-group structure can be used without any loss in risk detection.

Modifying the ratio weighting structure. As explained above, GMS uses a two-step weighting system to assign importance to the eight terms used to score banks. Since it seems unlikely that the importance of any bank activity or growth in detecting risk is stable over time, a periodic resetting of GMS term weights is necessary. The GMS User Manual does not state how the present weighting structure was chosen. In this section we present a means of determining GMS term weights on the basis of the importance of the eight growth measures in a model forecasting CAMEL rating downgrades. Specifically, the eight GMS terms were used as explanatory variables in a logit model relating the growth measures to the incidence of CAMEL rating downgrades occurring three years after growth was measured (see table 13.8).

Table 13.8

Relationship between GMS Weightings and Logit Estimations of CAMEL Downgrades

GMS Term	Initial Weight	Final Weight	December 1988 Model Coefficient
Asset growth	.60	0.67	0.0022 ns
Loan growth	.60	0	0.0053 *
Loans and sec/assets growth	.60	0.11	20.0043 *
Volatile liab/assets growth	.60	0.11	20.0016 ns
Equity/assets growth	.60	0.11	0.0005 ns
Loans and sec/assets ratio	.40	0.11	0.0111 *
Volatile liab/assets ratio	.40	0.11	0.0085 *
Equity/assets ratio	.40	0.11	0.0063 *

Note: The asterisk denotes significance at the 1 percent confidence level, and "ns" denotes "not significant" at the 1 percent confidence level.

The coefficient estimates from logit estimation of the CAMEL downgrade model were fairly consistent over time. Comparisons of the GMS weights and logit coefficients show that whereas GMS placed the greatest weight on asset growth and the least weight on loan growth, the logit model argued in favor of doing the reverse. Logit estimations indicated that whereas loan growth was significantly related to changes in condition (CAMEL downgrades), asset growth was not. In addition, the logit estimations argued for placing much greater emphasis (weight) upon the three ratios—loans and securities to assets, volatile liabilities to assets, and equity to assets—than do the present GMS weightings. The next subsection discusses tests made to see how GMS would be enhanced if the system were reweighted. It also discusses tests made with the use of additional risk measures.

Adding to the variables. The focus of GMS can be broadened to consider potentially risky changes in bank loan concentrations or shifts in business activity that may occur with (or without) growth in total loans or assets. For example, during the 1980s and 1990s the risks associated with rapid loan growth were often linked to increased portfolio concentrations in risky areas such as commercial real estate—a type of growth that is presumably riskier than growth in safe loan products such as residential mortgages. Yet GMS does not distinguish between these or other types of growth. To GMS, growth in residential mortgages is no different from growth in unsecured loans or loans with questionable collateral values.

Shifts in business activity can be measured with summary measures of loan portfolio concentration, analogous to the Herfindahl-Hirschman Index (HHI).[25] The bank portfolio concentration index proposed here is a summary measure of loan concentration for an individual bank. First, to measure overall loan concentration one computes the shares of total loans for several well-defined categories of loans. Next, to form the loan concentration index one squares and sums the portfolio shares.

Table 13.9 presents two hypothetical cases for Bank A. In Case 1, Bank A replaced $20 of residential mortgages with construction loans, increasing its loan portfolio concentration in the process. The portfolio concentration index increases from 2,500 to 3,300 (+32 percent) between 1980 and 1981 in Case 1. In Case 2, the bank reduced long-term commercial real estate by $20 and increased construction loans by the same amount, and the same concentration increase occurs. The portfolio shifts in Case 1 and Case 2, however, are not equal in terms of overall risk exposures. Most observers would agree that long-term commercial real estate and construction loans are riskier loan categories than residential

[25] The HHI is a measure of product market concentration and measures concentration of market shares across competitors in a market. The HHI is defined as the sum of squared market shares for all competitors in a well-defined geographic or product market. High HHI values indicate the concentration of market power (shares) among a few firms, while low HHI values indicate higher levels of competition.

Table 13.9
Hypothetical Loan Portfolios for Bank A: Loan Shares Not Weighted

		1981	
	1980	Case 1	Case 2
Residential mortgages	$25	$ 5	$25
Commercial and industrial	25	25	25
Commercial real estate	25	25	5
Construction and land development	25	45	45
Loan concentration index	2,500	3,300	3,300
Percent change in loan concentration (unweighted)		+32%	+32%

mortgages. Consequently Case 1, where total commercial real estate and construction loan exposures are higher, should be treated differently from Case 2. The way to distinguish these cases is by weighting loan portfolio shares, giving greater weight to riskier loan categories. A weighted portfolio concentration index would show greater increases in (weighted) concentration when overall risk exposures increase. Using the previous example but giving commercial real estate and construction loans greater weight in the portfolio concentration index, one finds that in Case 1 the overall risk exposure is now greater than in Case 2 (see table 13.10).

To test the usefulness of the concentration index, the researchers devised a loan portfolio concentration index by dividing total loans into 15 loan categories and weighting riskier loan shares more heavily than other loan categories. The loan categories and weights used are presented in table 13.11.

Table 13.10
Hypothetical Loan Portfolios for Bank A: Loan Shares Weighted

			1981	
	Weight	**1980**	Case 1	Case 2
Residential mortgages	0%	$25	$ 5	$25
Commercial and industrial	100	25	25	25
Commercial real estate	200	25	25	5
Construction and land development	200	25	45	45
Loan concentration index		3,125	5,925	4,700
Percent change in loan concentration (weighted)			+89.6%	+50%

Table 13.11
Loan Portfolio Concentration Index

	Weight
Loans Secured by Real Estate	
Construction loans	5
Secured by farmland	0
1- to 4-Family residential	0
Multifamily (5+) dwellings	5
Other commercial properties	5
Other real estate loans	0
Loans Not Secured by Real Estate	
Loans to banks	0
Agricultural loans	0
Commercial and industrial	1
Acceptances of U.S. banks	0
Consumer loans	1
Loans to foreign govts and orgs.	1
Municipal loans	0
Total leases	0
Unearned income	0

These weights were chosen somewhat arbitrarily to reflect the riskiness of commercial real estate loans relative to other loan areas. More-precise weights could be based upon loss experience by loan type.[26]

The researchers tested the relationships between revised GMS score rankings and changes in CAMEL ratings three years after scores were computed. The GMS scores were revised in two ways. First, the portfolio concentration index just described was included as a scoring variable, as was the percentage change in the portfolio concentration index over the growth quarter. Second, the percentile ranks of the ten terms in the revised GMS score (the original eight terms [table 13.8] plus portfolio concentration and growth in portfolio concentration) were weighted by use of the coefficients obtained from a logit model that related the ten terms to changes in CAMEL ratings, as discussed above. (See table 13.12.) To

[26] Concentration measures based upon income or revenue shares might also be better measures of risk. For example, the proportion of total revenue generated by each loan type might be used to form a concentration index normalized by revenue. Such measures will be tested in the future.

Table 13.12
12/1988 CAMEL Logit

GMS Term	Model Coefficient
Portfolio concentration	0.0025*
Growth in portfolio conc.	0.0060*
Asset growth	0.0022 ns
Loan growth	0.0054*
Loans and sec/assets growth	20.0044*
Volatile liab/assets growth	20.0017 ns
Equity/assets growth	0.0005 ns
Loans and sec/assets ratio	0.0111*
Volatile liab/assets ratio	0.0068
Equity/assets ratio	0.0067*

Note: The asterisk denotes significance at the 1 percent confidence level, and "ns" denotes "not significant" at the 1 percent level.

weight the terms, estimated coefficients from 1988 were chosen arbitrarily. Terms that were not statistically significant were given zero weight.

The analysis indicated that in each year the revised GMS score rankings were improved indicators of the likelihood of future CAMEL rating downgrades. For example, in December 1984 the proportion of downgraded banks in the highest (original) GMS decile was 15.96 percent (table 13.7), but with the revised GMS score the proportion increased to 18.45 percent. Similarly, with the revised GMS score the proportion of downgraded banks in the top two deciles increased from 28.07 percent to 32.4 percent (see table 13.13). In every year between 1984 and 1995 the highest revised GMS decile contained greater concentrations of downgraded banks than did the highest decile for the original GMS score. Even greater improvement can be expected to follow from less-arbitrary loan weightings as well as from adjustments in GMS term weightings over time.

The FDIC's Examination Ratings Model (CAEL)

Identifying factors that might affect bank performance several years in the future is not the same thing as identifying banks with deteriorating financial condition between examinations so that examination resources can be efficiently targeted at the identified institutions. Identifying the institutions with deteriorating condition has been the general focus of early-warning systems. And just as in a civil suit a preponderance of the evidence is required to reject the null hypothesis of innocence in favor of guilty, in a surveillance system for problem-bank detection the equivalent of a preponderance of the evidence is a high

Table 13.13
Comparisons of Exam Ratings as Assigned in 1985 and 1987
Portfolio Concentration Model
(Number and Column Percent)

December 1984 GMS Decile	CAMEL Upgraded		CAMEL Downgraded		No CAMEL Change		Missing Because of Mergers, etc.		Total	
1	203	16.29%	54	5.06%	384	11.47%	825	9.16%	1,466	9.99%
2	144	11.56	75	7.02	387	11.56	861	9.56	1,467	10.00
3	157	12.60	75	7.02	391	11.68	844	9.37	1,467	10.00
4	140	11.24	75	7.02	337	10.07	915	10.16	1,467	10.00
5	138	11.08	95	8.90	350	10.45	884	9.81	1,467	10.00
6	106	8.51	106	9.93	346	10.33	909	10.09	1,467	10.00
7	111	8.91	122	11.42	351	10.48	883	9.80	1,467	10.00
8	114	9.15	120	11.24	279	8.33	954	10.59	1,467	10.00
9	85	6.82	149	13.95	291	8.69	942	10.46	1,467	10.00
10	48	3.85	197	18.45	232	6.93	990	10.99	1,467	10.00
Total	**1,246**	**100.00%**	**1,068**	**100.00%**	**3,348**	**100.00%**	**9,007**	**100.00%**	**14,669**	**100.00%**

measured probability that a flagged institution will turn out to be a problem bank (so that few healthy banks are wrongly flagged as problem banks). Such systems should have a small Type II error.

Historical Development

In 1977, after the OCC developed the National Bank Surveillance System (NBSS), the FDIC introduced the Integrated Monitoring System (IMS). IMS consisted of 12 "tests": a set of Call Report ratios and associated benchmarks performed on commercial banks whose composite CAMEL ratings at their last examinations were 1 or 2. Institutions failing these benchmarks were flagged for the analytic attention of regional and field office personnel. Over time, the IMS failed to achieve a small Type II error. For example, in an IMS Failure Report for the first quarter of 1984, 2,758 institutions (more than 39 percent of the 7,400 institutions scored) were given "top-priority flags." Included were 100 of the 168 (60 percent) institutions having assets of $1 billion or more. The inclusion of so many high-asset institutions may have occurred for a number of reasons, but probably when the flag cutoffs were established the system did not take into account the differences in operation between banks of different sizes.

In response to these problems, the IMS off-site monitoring approach was given a major overhaul. This led to the FDIC's CAEL model that, like the IMS, was based solely on

Call Report data. The model was introduced at the end of the December 1985 Call Report processing period, and it remains the principal tool for the FDIC's off-site monitoring system. The CAEL model is named after the first letter of four of the five examination rating components: capital, asset quality, earnings, and liquidity (the "M," management, is not modeled). CAEL is used to perform a variety of tasks designed to help achieve and maintain efficient allocation of supervisory resources, primarily by early detection of banks that appear to have a high probability of a rating downgrade.

Description

CAEL is an "expert system," designed to replicate the financial analysis that an examiner would perform to assign an examination rating. As such, the system is a nonparametric, nonstatistical construct. CAEL is designed to predict examination ratings that would have been assigned if an institution had been examined as of the date of a Call Report. CAEL uses 19 financial ratios that are matched within peer groups.[27]

The FDIC periodically updates the CAEL model by having analysts subjectively determine new weights for each of the relevant CAEL ratio components. In addition, the CAEL component rating tables are updated each quarter to mirror the proportionate distribution of peer-group examination ratings over the previous year. Through this process, the CAEL rating distribution always approximates the previous year's examination rating distribution. The final component ratings are multiplied by their respective weights and combined to generate a single CAEL composite rating for an institution. CAEL component and composite ratings range from 0.50 (best) to 5.49 (worst), a range that corresponds to the examination CAMEL rating range of 1 to 5.

Banks' CAEL ratings are compared with their most recent composite CAMEL ratings. If the result is a large predicted downgrade from the current CAMEL rating, the appropriate Division of Supervision regional office gives those banks increased supervisory attention. Regional personnel are required to review these lists and, for each institution, determine whether they agree or disagree with CAEL's results. If they agree, appropriate supervisory follow-up of the subject institution must take place (appropriateness is a function of the severity of the CAEL rating and the gap between the CAEL rating and the CAMEL rating). If regional staff disagree with CAEL's results, the reason(s) for the disagreement must be documented in writing and transmitted to the Washington, D.C., office. In volume the CAEL Off-Site Review List has ranged from 1 percent to 3 percent of the institutions modeled.[28] In addition to following up on the Off-Site Review List, regional staff use CAEL to help in such

[27] There are three commercial-bank peer groups based solely on asset size; a fourth group encompasses all FDIC-supervised savings institutions.

[28] During the study period (1987–94), CAEL encompassed only about 80 percent of all commercial banks. It excluded all banks with over $1 billion in assets and all banks with a CAMEL rating of 1.

supervisory activities as scheduling exams, doing preexamination planning, assessing affiliated holding company institutions and chain banking organizations,[29] and reviewing risk-related premium classifications.

Validation

CAEL has been in existence since 1985, but only since 1987 have sufficient examinations been conducted for there to be measurable results. The validation analysis, for the period from 1987 to 1994, focuses on how well the system performs its stated functions, which are to identify deteriorating banks so that exam resources can be efficiently targeted and to do so with a high probability that a flagged institution will in fact have deteriorated (high probability of guilt).

From 1987 to 1994, only 16 percent of banks with composite CAMEL ratings of 2 or lower experienced a rating downgrade. If there had been no off-site system and the bank regulators had simply randomly chosen banks for accelerated examinations, 16 percent of the examinations would have resulted in a rating downgrade. Over the same period, 52 percent of all CAEL-predicted downgrades were actually downgraded within six months. In other words, CAEL was more than three times better than a random draw at predicting downgrades. Note that a random draw would result in a Type I error rate of 84 percent (84 percent of institutions examined would not be downgraded), while the CAEL model has a Type I error rate of 48 percent, a very large improvement.

Another way to analyze CAEL's effectiveness is to see what percentage of total downgrades were identified. During the period there were 2,867 downgrades, 715 of which—or 25 percent of the relevant group—CAEL predicted. Although CAEL does not predict downgrades for CAMEL 1-rated institutions, some of these institutions were downgraded, and a large number by more than one rating, that is, CAMEL 1 to CAMEL 3. If these downgrades are included, CAEL correctly predicted only 14.3 percent of the total of 3,810 downgrades. By design, CAEL will miss a large number of actual downgrades in order to avoid targeting banks that are, in fact, in sound condition.

The Federal Reserve Board's Financial Institutions Monitoring System (FIMS)

During the years when the OCC and the FDIC were developing their first off-site systems, the Federal Reserve Board developed a similar system,[30] a system of screens, which it replaced in the mid-1980s with the Uniform Bank Surveillance System (UBSS), an out-

[29] "Chain banking organizations" refer to banks that are controlled by the same ownership group but are not associated with a bank holding company.

[30] Putnam, "Early-Warning Systems."

growth of the OCC's early National Bank Surveillance System (NBSS).[31] After the UBSS and CAEL were developed, a substantial body of economic research focused on modeling bank failures and financial distress.[32] This research indicated that banks' financial condition could be successfully modeled with the use of standard Call Report data and that the models would probably use far fewer variables than CAEL. Taking note of the research, in 1993 the FRB replaced the UBSS with the Financial Institutions Monitoring System (FIMS).[33]

FIMS represented a major advance in surveillance systems by using sophisticated statistical models to predict CAMEL ratings (ordinal-level logit) and probabilities of commercial bank failures (binary-probit). These techniques allow the bank analyst to determine statistically what bank condition ratios are significant determinants of CAMEL ratings or of failure, and how important each ratio might be in the model. The techniques also help the analyst discard ratios that do not have a statistically significant relationship with CAMEL ratings or bank failures. These models can also be updated (reestimated) as often as four times a year (when new Call Report data are received) to adapt to changes in examination standards or in the banking environment.

The FIMS model lends itself to the same type of validation that was performed for CAEL. The FIMS validation covers the period from December 1989 to March 1992, a much shorter period than that used for validating the CAEL system.[34] Over its validation period FIMS correctly identified 61 percent of downgrades predicted, with a 39 percent Type II error rate. FIMS also identified 41.2 percent of the total downgrades. For purposes of comparison, a CAEL validation was performed for the same period. CAEL correctly identified 51 percent of downgrades (10 percent below FIMS), with a 49 percent Type II error rate. CAEL also predicted a smaller percentage of total downgrades than FIMS, 34 percent (41.2 percent for FIMS). Thus, over the same study period, FIMS was more accurate than CAEL.

It should be noted that there are meaningful differences between an "algorithmic" model like CAEL and a "probabilistic" model like FIMS. FIMS estimates the "probability" of bank failure or of rating downgrade by using historical trends and relationships. CAEL is not based on any statistical model and does not require assumptions about standard statistical problems, such as normality in dependent-term distribution. But CAEL's nonpara-

---

[31] Rebel A. Cole, Barbara G. Cornyn, and Jeffery W. Gunther, "FIMS: A New Monitoring System for Banking Institutions," *Federal Reserve Bulletin* 81 (January 1995): 3.

[32] Alst Demirguc-Kurt, "Modeling Large Commercial-Bank Failures: A Simultaneous-Equations Analysis," working paper 8905, Federal Reserve Bank of Cleveland, May 1989; Alst Demirguc-Kurt, "Deposit-Institution Failures: A Review of the Empirical Literature," Federal Reserve Bank of Cleveland *Economic Review* 27 (March 1991); Gregory R. Gajewski, "Assessing the Risk of Bank Failure," in *Bank Structure and Competition,* conference proceedings, Federal Reserve Bank of Chicago (May 1989), 432—56; and Gary Whalen and James Thomson, "Using Financial Data to Identify Changes in Bank Conditions," Federal Reserve Bank of Cleveland *Economic Review* (quarter 2, 1988): 17–26.

[33] Cole, Cornyn, and Gunther. "FIMS: A New Monitoring System," 1–15.

[34] Ibid.

metric approach is nonetheless dependent upon historical patterns and is therefore still open to the same weaknesses as probabilistic models. In addition, the Type I and Type II error trade-offs are inherent in all forecasts regardless of estimation method.

After FIMS, the FRB went on to make many improvements to its entire surveillance system. The bank-failure model described in FIMS is no longer in use; the model was reestimated quarterly using the previous two years' data. But as the number of bank failures dramatically decreased through the 1990s, the model became less and less reliable.[35] The FRB then developed a new failure model: a pooled time series model that uses failures for the years 1985 through 1993 to estimate failure probabilities. Building on the original FIMS work, researchers studied a large number of aggregate economic variables (Treasury bond rates, changes in GDP, etc.) to see if they increase the predictive power of either the CAMEL ratings model or the failure model. Although some of the variables were statistically significant, none of them improved the accuracy of the "out-of-sample" predictions of the models. The FRB also continues to develop and use various financial-ratio screens that highlight outlier banks. Currently the ratings model is updated quarterly, and the pooled failed-bank model is updated every two years. All parts of the FRB surveillance system are produced and distributed within three days of final Call Report data.

The OCC's Surveillance System

The OCC's current off-site surveillance system uses a variety of mainframe and PC applications based on Call Report and UBPR information. Two of the PC systems use artificial intelligence and expert system technology. One of the two takes the UBPR for each national and state-chartered bank and produces an English-language report based on expert financial analysts' experience. The other analyzes the interest-rate risk of each bank based on historical changes, and produces an English-language summary of the findings. Summary scores from both systems provide for trend and systematic analysis across all banks. For national banks, these reports are produced within a day after quarterly Call Report data are final.

Conclusion

The lessons of the banking crises of the 1980s and the use of off-site monitoring during that period are fairly clear. First, banks that either become financially distressed or fail apparently exhibit identifiable risk characteristics several years in advance of the distress or failure. Second, off-site surveillance systems like those now used by the FRB appear to be reliable and valuable tools when used in conjunction with regular examinations. Finally, ongoing research at each of the federal bank regulatory agencies is warranted. Such research would include regional economic data in the current off-site monitoring models, thereby further enhancing our understanding of the causes of financial distress.

---

[35] The model did not have enough failure "events" for the statistical procedure used to run to a solution.

Appendix A

The tables in this appendix present a comparison of different factors in predicting bank failures four and five years into the future for the years 1980, 1982, 1984, 1986, and 1988. The banks studied consisted of all banks that were in existence on December 31 of the beginning year (for instance 1980) and either failed four or five years later or never failed. Excluded from the analysis were banks that failed before the fourth or fifth year, banks that were chartered after the beginning period, and banks that failed subsequent to the fifth year.

There is a table for each beginning year of the study, and three groups of banks are shown within each table. The first part of each table contains data from the entire set of banks studied for the particular time period. This analysis (described below) identifies the subset of banks that exhibit the highest risk of bank failure within the universe of banks studied. The second part of the table then repeats the analysis using only the highest-risk subset of banks. The third part of the table takes the remaining banks in the universe (those not identified as the highest risk in the first part of the table) and again repeats the analysis.

Table 13-A.1 compares the different risk factors as predictors of bank failures four and five years forward from 1980. The eight financial ratios that were chosen as risk factors for the analysis are listed across the top of the table. For each ratio, banks are sorted from lowest to highest and divided into five equal portions or quintiles, with banks in quintile 1 supposedly having the lowest risk and those in quintile 5 the highest. By reading across each quintile, one can identify how many banks failed and how many banks never failed for each risk factor. For instance, in the first quintile for Loans to Assets for 1980 (the 20 percent of the banks with the lowest Loans-to-Assets ratio), 20 banks failed either in 1984 or 1985 and 2,415 never failed. The first quintile for Return on Assets (ROA) in 1980 had 39 failed banks four or five years later and 2,396 banks that never failed. The Total row shows that 184 of the banks in existence in 1980 failed either in 1984 or 1985 and that 11,989 banks in existence in 1980 never failed.

Below the Total row are the Chi-Square statistics for the logit regression for each risk factor. The higher the Chi-Square statistic, the better the risk factor is as a predictor of bank failures. In 1980 the Loans-to-Assets ratio is the risk factor with the highest Chi-Square: 99.668. The risk factor with the next-highest score is Loan Growth, with a Chi-Square of 88.352.

The last row is labeled Best Grouping and identifies which set of banks within the high-risk ratio is the best predictor of failure. Obviously banks in quintile 5 had the highest failure rate (88 out of 2,434), but it may be that the best predictor of failure was being a bank in either quintile 4 or quintile 5. To determine which grouping of banks yields the best prediction of failure, four groupings are analyzed: quintile 1 versus quintiles 2–5, quintiles 1–2 versus quintiles 3–5, and so forth. A Chi-Square statistic is calculated on the difference between each two groups. The largest Chi-Square indicates the largest difference between the groups, and being in the group with the highest percentage of failures is the best predictor of failure. For the Loans-to-Assets factor, the smallest difference was between quintile 1 and quintiles 2–5, with a Chi-Square of 9.74. The largest difference was between the grouping of quintiles 1–4 versus quintile 5, with a Chi-Square of 90.4. Of the 20 percent of the banks in quintile 5, 3.6 percent failed, over three times the failure rate (1 percent) of the other 80 percent of the banks. Thus, in 1980 a bank with a Loans-to-Assets ratio in the highest 20 percent of all banks was the best predictor of failure in 1984 or 1985.

In the next part of the table, banks in the highest-risk grouping (that is, quintile 5 of Loans to Assets) are analyzed to determine if there were additional risk factors that were related significantly to failure. Although the 2,434 highest-risk banks identified so far were all in the top 20 percent of Loans to Assets, they are not uniformly distributed in the highest-risk quintiles of the other risk factors. For example, the 2,434 institutions (88 failures and 2,346 nonfailures) are spread relatively evenly through the Average Salary quintiles but very unevenly through the Interest Yield quintiles, where a large percentage of banks were concentrated in quintiles 4 and 5. Interest Yield would appear to be an excellent predictor of risk, as 78 of the 88 failures (89 percent) are in the fourth and fifth quintiles. However, a very large percentage of all of the high-risk banks (65.3) are in those quintiles, so the prediction may not be as good as it first appears. In fact, the Chi-Square for the best predictor for the Interest Yield is 36.05, the second-highest for the High-Risk Group. The best predictor turned out to be Interest and Fees on Loans and Leases (Chi-Square of 52.22). For this predictor, the best grouping was between quintiles 1–3 and quintiles 4–5. A bank in both the top 20 percent of Loans to Assets and the top 40 percent of Interest and Fees to Loans and Leases would have a 7.2 percent probability of failure, twice the rate of being only in the top quintile of Loans to Assets.

The final part of the table is the risk analysis of the banks that were in the lower 80 percent of Loans to Assets. The same procedure that was used for the high-risk banks was followed, and the results are shown in the Low-Risk Group section of table 13-A.1. The variable with the highest Chi-Square statistic (42.61) is Loan Growth. The best grouping was between quintiles 1–4 and quintile 5 (Chi-Square of 36.26). In quintile 5, 2.3 percent of banks failed, versus 0.7 percent of the remaining banks.

To summarize, on December 31, 1980 there were 12,173 banks that either failed in 1984 or 1985 or never failed. Of that group, 184 failed four or five years later. Of the risk factors studied, the banks with the highest probability of failure were those in the highest quintile for Loans to Assets and in the highest 40 percent of banks for Interest and Fees on Loans and Leases. For the 80 percent of banks that were not in the High-Risk Group, being in the highest Loan Growth quintile was the best predictor of failure.

Table 13-A.1

Comparison of Different Factors in Predicting Bank Failures Four and Five Years Forward, 1980

	Loans to Assets		ROA		Asset Growth		Operating Expenses		Interest Yield		Average Salary		Int and Fees to Loans		Loan Growth	
Quintile	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed
All Banks																
1	20	2,415	39	2,396	36	2,398	25	2,409	23	2,412	45	2,390	34	2,400	21	2,414
2	13	2,421	31	2,404	16	2,419	20	2,414	17	2,418	19	2,416	14	2,421	13	2,421
3	25	2,410	24	2,410	32	2,403	36	2,399	18	2,416	38	2,397	30	2,405	29	2,406
4	38	2,397	36	2,399	34	2,400	36	2,399	44	2,390	30	2,403	52	2,383	36	2,399
5	**88**	**2,346***	54	2,380	66	2,369	67	2,368	82	2,353	52	2,383	54	2,380	85	2,349
Total	184	11,989	184	11,989	184	11,989	184	11,989	184	11,989	184	11,989	184	11,989	184	11,989
Chi-Square	**99.668**		13.769		36.323		36.803		83.604		18.274		30.382		88.352	
Best Grouping	(1 vs 2–5) 9.740				(1–2 vs 3–5) 37.895				(1–3 vs 4–5) 63.140				(1–4 vs 5) **90.404**			
High-Risk Group																
1	0	0	28	356	20	557	11	449	3	135	16	435	7	657	3	158
2	0	0	15	374	9	560	11	401	4	279	9	490	5	489	6	345
3	0	0	11	462	14	476	14	452	3	438	17	488	18	453	15	517
4	0	0	16	486	18	410	18	497	19	620	13	478	**30**	**406***	19	660
5	88	2,346	18	668	27	343	34	547	59	874	33	455	**28**	**341***	45	666
Total	88	2,346	88	2,346	88	2,346	88	2,346	88	2,346	88	2,346	88	2,346	88	2,346
Chi-Square			19.415		22.420		11.896		36.057		19.984		**52.216**		22.371	
Best Grouping	(1 vs 2–5) 17.189				(1–2 vs 3–5) 42.167				(1–3 vs 4–5) **44.473**				(1–4 vs 5) 19.681			
Low-Risk Group																
1	20	2,415	11	2,040	16	1,841	14	1,960	20	2,277	29	1,955	27	1,743	18	2,256
2	13	2,421	16	2,030	7	1,859	9	2,013	13	2,139	10	1,926	9	1,932	7	2,076
3	25	2,410	13	1,948	18	1,927	22	1,947	15	1,978	21	1,909	12	1,952	14	1,889
4	38	2,397	20	1,913	16	1,990	18	1,902	25	1,770	17	1,925	22	1,977	17	1,739
5	0	0	36	1,712	39	2,026	33	1,821	23	1,479	19	1,928	26	2,039	**40**	**1,683***
Total	96	9,643	96	9,643	96	9,643	96	9,643	96	9,643	96	9,643	96	9,643	96	9,643
Chi-Square			27.905		25.469		19.978		12.260		9.423		15.379		**42.611**	
Best Grouping	(1 vs 2–5) 1.146				(1–2 vs 3–5) 13.71				(1–3 vs 4–5) 23.62				(1–4 vs 5) **38.26**			

*Quintile that is best predictor of failure.

Table 13-A.2

Comparison of Different Factors in Predicting Bank Failures Four and Five Years Forward, 1982

	Loans to Assets		ROA		Asset Growth		Operating Expenses		Interest Yield		Average Salary		Int and Fees to Loans		Loan Growth	
Quintile	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed
								All Banks								
1	8	2,363	54	2,318	63	2,306	37	2,333	24	2,345	54	2,317	29	2,340	25	2,345
2	25	2,347	55	2,316	47	2,325	49	2,322	20	2,353	45	2,326	27	2,345	23	2,348
3	36	2,333	50	2,323	40	2,331	45	2,326	33	2,338	45	2,328	52	2,318	33	2,339
4	62	2,311	53	2,319	56	2,317	62	2,311	58	2,314	63	2,307	76	2,296	88	2,285
5	**160**	**2,212***	79	2,290	85	2,287	98	2,274	156	2,216	84	2,288	107	2,267	122	2,249
Total	291	11,566	291	11,566	291	11,566	291	11,566	291	11,566	291	11,566	291	11,566	291	11,566
Chi-Square	**255.143**		9.838		21.183		40.572		225.857		18.58		80.164		139.722	
Best Grouping		(1 vs 2–5) 55.472				(1–2 vs 3–5) 102.109				(1–3 vs 4–5) 163.495				(1–4 vs 5) **227.845**		
								High-Risk Group								
1	0	0	31	359	33	582	18	381	7	161	26	391	9	524	9	195
2	0	0	29	369	26	445	21	333	7	275	25	439	18	458	13	331
3	0	0	32	391	19	398	28	393	13	376	24	501	30	443	14	436
4	0	0	24	395	39	352	42	460	30	542	39	436	**50**	**456***	55	585
5	160	2,212	44	698	43	435	51	645	103	858	46	445	**53**	**331***	69	665
Total	160	2,212	160	2,212	160	2,212	160	2,212	160	2,212	160	2,212	160	2,212	160	2,212
Chi-Square			3.009		16.491		6.016		43.245		12.469		**66.747**		27.729	
Best Grouping		(1 vs 2–5) 27.947				(1–2 vs 3–5) 46.228				(1–3 vs 4–5) **52.778**				(1–4 vs 5) 36.207		
								Low-Risk Group								
1	8	2,363	23	1,724	30	1,724	19	1,952	17	2,184	28	1926	20	1,816	16	2,150
2	25	2,347	26	1,880	21	1,880	28	1,989	13	2,078	20	1887	9	1,887	10	2,017
3	36	2,333	18	1,933	21	1,933	17	1,933	20	1,962	21	1827	22	1,875	19	1,903
4	62	2,311	29	1,965	17	1,965	20	1,851	28	1,772	24	1871	26	1,840	33	1,700
5	0	0	35	1,852	42	1,852	47	1,629	**53**	**1,358***	38	1843	54	1,936	53	1,584
Total	131	9,354	131	9,354	131	9,354	131	9,354	131	9,354	131	9354	131	9,354	131	9,354
Chi-Square			11.009		17.524		32.494		**75.695**		8.221		39.209		65.372	
Best Grouping		(1 vs 2–5) 7.7978				(1–2 vs 3–5) 26.783				(1–3 vs 4–5) 46.436				(1–4 vs 5) **68.628**		

*Quintile that is best predictor of failure.

Table 13-A.3

Comparison of Different Factors in Predicting Bank Failures Four and Five Years Forward, 1984

	Loans to Assets		ROA		Asset Growth		Operating Expenses		Interest Yield		Average Salary		Int and Fees to Loans		Loan Growth	
Quintile	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed
								All Banks								
1	11	2287	44	2,255	90	2,200	34	2,260	42	2,254	45	2,250	61	2,233	64	2,227
2	22	2276	34	2,263	36	2,260	33	2,264	34	2,263	55	2,241	53	2,244	39	2,256
3	47	2249	42	2,257	35	2,264	60	2,239	47	2,251	62	2,233	58	2,240	44	2,255
4	64	2231	75	2,222	39	2,258	88	2,209	71	2,222	84	2,210	66	2,229	51	2,246
5	**188**	**2104***	137	2,150	132	2,165	117	2,175	138	2,157	86	2,213	94	2,201	134	2,163
Total	332	11147	332	11,147	332	11,147	332	11,147	332	11,147	332	11,147	332	11,147	332	11,147
Chi-Square	**314.339**		112.728		116.781		81.377		111.339		20.134		16.192		93.997	
Best Grouping		(1 vs 2–5) 59.593				(1–2 vs 3–5) 129.011				(1–3 vs 4–5) 184.104				(1–4 vs 5) **287.196**		
								High-Risk Group								
1	0	0	26	298	42	376	16	302	21	230	24	405	34	522	23	206
2	0	0	19	351	14	317	20	323	13	249	31	441	33	414	15	313
3	0	0	25	417	17	369	35	366	23	324	35	401	38	423	24	411
4	0	0	46	466	22	399	49	476	47	480	45	428	42	410	27	505
5	188	2,104	72	572	**93**	**643***	68	637	84	821	53	429	41	335	99	669
Total	188	2,104	188	2,104	188	2,104	188	2,104	188	2,104	188	2,104	188	2,104	188	2,104
Chi-Square			16.442		**40.392**		9.797		6.564		11.643		7.975		40.262	
Best Grouping		(1 vs 2–5) 2.312				(1–2 vs 3–5) 0.778				(1-3 vs 4–5) 9.363				(1–4 vs 5) **28.061**		
								Low-Risk Group								
1	11	2,287	18	1,957	48	1,824	18	1,958	21	2,024	21	1,845	27	1,711	41	2,021
2	22	2,276	15	1,912	22	1,943	13	1,941	21	2,014	24	1,800	20	1,830	24	1,943
3	47	2,249	17	1,840	18	1,895	25	1,873	24	1,927	27	1,832	20	1,817	20	1,844
4	64	2,231	29	1,756	17	1,859	39	1,733	24	1,742	39	1,782	24	1,819	24	1,741
5	0	0	**65**	**1,578***	39	1,522	49	1,538	54	1,336	33	1,784	53	1,866	35	1,494
Total	144	9,043	144	9,043	144	9,043	144	9,043	144	9,043	144	9,043	144	9,043	144	9,043
Chi-Square			**79.204**		33.559		44.975		57.976		7.892		24.152		12.514	
Best Grouping		(1 vs 2–5) **73.962**				(1–2 vs 3–5) 22.899				(1–3 vs 4-5) 48.909				(1–4 vs 5) **73.962**		

*Quintile that is best predictor of failure.

Table 13-A.4

Comparison of Different Factors in Predicting Bank Failures Four and Five Years Forward, 1986

	Loans to Assets		ROA		Asset Growth		Operating Expenses		Interest Yield		Average Salary		Int and Fees to Loans		Loan Growth	
Quintile	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed
All Banks																
1	7	2,247	18	2,237	66	2,171	32	2,220	53	2,195	40	2,208	72	2,175	47	2,190
2	14	2,239	25	2,229	31	2,219	25	2,228	28	2,222	33	2,216	50	2,200	33	2,214
3	26	2,221	29	2,225	27	2,225	29	2,219	38	2,216	55	2,191	43	2,208	40	2,213
4	61	2,187	50	2,202	28	2,225	47	2,204	46	2,204	43	2,204	45	2,202	43	2,211
5	**145**	**2,099***	131	2,100	101	2,153	120	2,122	88	2,156	82	2,174	43	2,208	90	2,165
Total	253	10,993	253	10,993	253	10,993	253	10,993	253	10,993	253	10,993	253	10,993	253	10,993
Chi-Square	**261.134**		177.554		85.494		128.176		42.67		29.761		12.301		41.059	
Best Grouping		(1 vs 2–5) 48.202				(1–2 vs 3–5) 108.822				(1–3 vs 4–5) 185.640				(1–4 vs 5) **225.957**		
High-Risk Group																
1	0	0	10	311	38	360	16	299	22	305	21	421	50	701	23	158
2	0	0	18	409	15	273	14	295	11	306	21	417	30	472	17	223
3	0	0	18	440	15	296	21	374	30	318	29	450	24	377	22	364
4	0	0	30	433	14	470	26	470	26	410	23	402	23	280	26	559
5	145	2,099	**69**	**506***	63	700	68	661	56	760	51	409	18	269	57	795
Total	145	2,099	145	2,099	145	2,099	145	2,099	145	2,099	145	2,099	145	2,099	145	2,099
Chi-Square			**43.549**		22.670		14.888		7.813		21.342		1.051		16.217	
Best Grouping		(1 vs 2–5) 6.940				(1–2 vs 3–5) 13.717				(1–3 vs 4–5) 30.233				(1–4 vs 5) **39.067**		
Low-Risk Group																
1	7	2247	8	1,926	28	1,811	16	1,921	31	1,890	19	1,787	22	1,474	24	2,032
2	14	2239	7	1,820	16	1,946	11	1,933	17	1,916	12	1,799	20	1,728	16	1,991
3	26	2221	11	1,785	12	1,929	8	1,845	8	1,898	26	1,741	19	1,831	18	1,849
4	61	2187	20	1,769	14	1,755	21	1,734	20	1,794	20	1,802	22	1,922	17	1,652
5	0	0	**62**	**1,594***	38	1,453	52	1,461	32	1,396	31	1,765	25	1,939	33	1,370
Total	108	8894	108	8,894	108	8,894	108	8,894	108	8,894	108	8,894	108	8,894	108	8,894
Chi-Square			**116.120**		34.973		81.976		27.438		10.198		1.602		19.816	
Best Grouping		(1 vs 2–5) 12.841				(1–2 vs 3–5) 34.958				(1–3 vs 4–5) 65.614				(1–4 vs 5) **110.775**		

*Quintile that is best predictor of failure.

Table 13-A.5

Comparison of Different Factors in Predicting Bank Failures Four and Five Years Forward, 1988

	Loans to Assets		ROA		Asset Growth		Operating Expenses		Interest Yield		Average Salary		Int and Fees to Loans		Loan Growth	
Quintile	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed	Failed	Nonfailed
								All Banks								
1	6	2135	16	2,127	31	2,109	30	2,111	19	2,123	18	2,118	31	2,110	27	2,112
2	8	2134	13	2,129	9	2,133	12	2,129	14	2,128	14	2,124	15	2,126	15	2,127
3	17	2123	21	2,121	16	2,125	17	2,125	14	2,127	16	2,122	27	2,114	14	2,127
4	30	2111	18	2,123	12	2,130	21	2,120	26	2,115	27	2,112	25	2,116	14	2,128
5	**72**	**2071***	65	2,074	65	2,077	53	2,089	60	2,081	58	2,098	35	2,108	63	2,080
Total	133	10574	133	10,574	133	10,574	133	10,574	133	10,574	133	10,574	133	10,574	133	10,574
Chi-Square	**111.600**		71.628		81.030		39.765		56.791		49.897		8.632		67.567	
Best Grouping		(1 vs 2–5)	20.186			(1–2 vs 3–5)	48.751			(1–3 vs 4–5)	75.492			(1–4 vs 5)	**97.888**	
								High-Risk Group								
1	0	0	8	396	14	203	22	412	11	162	7	430	18	532	9	145
2	0	0	10	408	4	271	7	341	6	229	7	405	11	509	8	298
3	0	0	15	394	7	350	10	358	6	350	7	405	15	379	9	423
4	0	0	9	414	8	528	10	437	14	527	13	377	10	345	9	564
5	72	2,071	30	459	39	719	23	523	35	803	**38**	**454***	18	306	37	641
Total	72	2,071	72	2,071	72	2,071	72	2,071	72	2,071	72	2,071	72	2,071	72	2,071
Chi-Square			17.269		24.812		9.279		11.051		**40.015**		7.869		20.454	
Best Grouping		(1 vs 2–5)	5.225			(1–2 vs 3–5)	12.674			(1–3 vs 4–5)	27.092			(1–4 vs 5)	**37.388**	
								Low-Risk Group								
1	6	2,135	8	1,731	17	1,906	8	1,699	8	1,961	11	1,688	13	1,578	18	1,967
2	8	2,134	3	1,721	5	1,862	5	1,788	8	1,899	7	1,719	4	1,617	7	1,829
3	17	2,123	6	1,727	9	1,775	7	1,767	8	1,777	9	1,717	12	1,735	5	1,704
4	30	2,111	9	1,709	4	1,602	11	1,683	12	1,588	14	1,735	15	1,771	5	1,564
5	0	0	**35**	**1,615***	26	1,358	30	1,566	25	1,278	20	1,644	17	1,802	26	1,439
Total	61	8,503	61	8,503	61	8,503	61	8,503	61	8,503	61	8,503	61	8,503	61	8,503
Chi-Square			**59.088**		38.601		39.578		33.524		9.119		6.912		34.985	
Best Grouping		(1 vs 2–5)	1.963			(1–2 vs 3–5)	12.803			(1–3 vs 4–5)	27.707			(1–4 vs 5)	**57.352**	

*Quintile that is best predictor of failure.

Appendix B

Table 13-A.6

UBPR Peer-Group Characteristics

Peer Group	Average Assets for Latest Quarter	Number of Banking Offices	Location
1	In excess of $10 billion	–	–
2	Between $3 billion and $10 billion	–	–
3	Between $1 billion and $3 billion	–	–
4	Between $500 million and $1 billion	–	–
5	Between $300 million and $500 million	3 or more	–
6	Between $300 million and $500 million	2 or fewer	–
7	Between $100 million and $300 million	3 or more	Metropolitan area
8	Between $100 million and $300 million	3 or more	Nonmetropolitan area
9	Between $100 million and $300 million	2 or fewer	Metropolitan area
10	Between $100 million and $300 million	2 or fewer	Nonmetropolitan area
11	Between $50 million and $100 million	3 or more	Metropolitan area
12	Between $50 million and $100 million	3 or more	Nonmetropolitan area
13	Between $50 million and $100 million	2 or fewer	Metropolitan area
14	Between $50 million and $100 million	2 or fewer	Nonmetropolitan area
15	Between $25 million and $50 million	2 or more	Metropolitan area
16	Between $25 million and $50 million	2 or more	Nonmetropolitan area
17	Between $25 million and $50 million	1	Metropolitan area
18	Between $25 million and $50 million	1	Nonmetropolitan area
19	Between $10 million and $25 million	2 or more	Metropolitan area
20	Between $10 million and $25 million	2 or more	Nonmetropolitan area
21	Between $10 million and $25 million	1	Metropolitan area
22	Between $10 million and $25 million	1	Nonmetropolitan area
23	Less than or equal to $10 million	–	Metropolitan area
24	Less than or equal to $10 million	–	Nonmetropolitan area
25	Were established within the last three years, and have assets less than or equal to $25 million		