Predicting Bank Failures in the 1980s

by James B. Thomson

James B. Thomson is an assistant vice president and economist at the Federal Reserve Bank of Cleveland. The author would like to thank David Altig, Brian Cromwell, and Ramon DeGennaro for helpful comments, and Lynn Seballos for excellent research assistance.

Introduction

From 1940 through the 1970s, few U.S. banks failed. The past decade was a different matter, however, as bank failures reached record post-Depression rates. More than 200 banks closed their doors each year from 1987 through 1989, while 1990 saw 169 banks fold. And because more than 8 percent of all banks are currently classified as problem institutions by bank regulators, failures are expected to exceed 150 per year for the next several years.[1] The recent difficulties in the commercial real estate industry, especially in the Northeast and the Southwest, will likely add to the number of problem and failed banks in the 1990s.

The increase in bank failures in the 1980s was accompanied by an increase in the cost of resolving those failures. Furthermore, the cost of failure per dollar of failed-bank assets, which is already high, may continue to rise. For banks failing in 1985 and 1986, failure resolution cost estimates averaged 33 percent of failed-bank assets, while the estimated loss to the Federal Deposit Insurance Corporation (FDIC) reached as high as 64 percent of bank assets (see Bovenzi and Murton [1988]).

One characteristic that is different for some of the recent failures is bank size, as large-bank failures became more common in the 1980s. In 1984, for example, the FDIC committed $4.5 billion to rescue the Continental Illinois National Bank and Trust Company of Chicago, which at that time had $33.6 billion in assets. In 1987, BancTexas and First City Bancorporation of Dallas were bailed out by the FDIC at a cost of $150 million and $970 million, respectively. The $32.5 billion-asset First Republic Bancorp of Dallas collapsed in 1988, costing the FDIC approximately $4 billion, while 20 bank subsidiaries of MCorp of Houston, with a total of $15.6 billion in assets, were taken over by the FDIC in 1989 at an estimated cost of $2 billion.[2] Most recently, the Bank of New England, with $22 billion in assets, was rescued by the FDIC at an estimated cost of $2.3 billion.

■ 1 Examiners rate banks by assessing five areas of risk: capital adequacy, asset quality, management, earnings, and liquidity. This is called the CAMEL rating. For an in-depth discussion of the CAMEL rating system, see Whalen and Thomson (1988).

■ 2 In the 1980s, other large banks such as Texas American Bankshares, National Bank of Texas, First Oklahoma, and National of Oklahoma were either merged or sold with FDIC assistance. In addition, Seafirst of Seattle, Texas Commerce Bankshares, and Allied Bankshares had to seek merger partners to stave off insolvency.

The study of bank failures is interesting for two reasons. First, an understanding of the factors related to an institution's failure will enable us to manage and regulate banks more efficiently. Second, the ability to differentiate between sound banks and troubled ones will reduce the expected cost of bank failures. In other words, if examiners can detect problems early enough, regulatory actions can be taken either to prevent a bank from failing or to minimize the cost to the FDIC and thus to taxpayers. The ability to detect a deterioration in bank condition from accounting data will reduce the cost of monitoring banks by lessening the need for on-site examinations (see Benston et al. [1986, chapter 10] and Whalen and Thomson [1988]).

An extensive literature on bank failures exists.[3] Statistical techniques used to predict or to classify failed banks include multivariate discriminate analysis (Sinkey [1975]), factor analysis and logit regression (West [1985]), event-history analysis (Lane, Looney, and Wansley [1986, 1987] and Whalen [1991]), and a two-step logit regression procedure suggested by Maddala (1986) to classify banks as failed and nonfailed (Gajewski [1990] and Thomson [1989]). Recently, Demirgüç-Kunt (1989a, 1991, and forthcoming) has extended this work to include market data and a model of the failure decision. Unfortunately, market data are available only for the largest banking institutions, while the majority of banks that fail are small.

This study uses 1983–1988 book data from the June and December Federal Financial Institutions Examination Council's Reports of Condition and Income (call reports) in statistical models of bank failure. In addition to traditional balance-sheet and income-statement measures of risk, the failure equation incorporates measures of local economic conditions.

The historically high number of failures for every year in the sample period allows each year to be investigated separately. Previous studies had to pool the failures across years to obtain a sufficiently large failed-bank sample, making it difficult to construct holdout samples and to do out-of-sample forecasting. This was especially true for tests across years. The sample in this study is not limited in this way, however. Once failures for a particular year are classified by the model, failures in subsequent years can be used to determine the model's out-of-sample predictive ability. For example, the failure prediction model used to classify failures in 1985 can be applied to the 1984 data for banks that failed in 1986 and 1987.

I. Modeling Bank Failures

The economic failure of a bank occurs when it becomes insolvent. The official failure of a bank occurs when a bank regulator declares that the institution is no longer viable and closes it.[4] Insolvency is a necessary condition for regulators to close a bank, but not, Kane (1986) argues, a sufficient one. He suggests that the FDIC faces a set of four constraints on its ability to close insolvent banks. These constraints, which arise because of imperfect information, budget limitations, and principal–agent conflicts, include information constraints, legal and political constraints, implicit and explicit funding constraints, and administrative and staff constraints (see Kane [1989]). Both Thomson (1989, 1991) and Demirgüç-Kunt (1991) formally incorporate Kane's constraints on the FDIC's ability to close banks into models of the closure decision. These authors, along with Gajewski (1990), estimate two-equation models that formally separate economic insolvency and closure.

The model in this paper is a variant of those in the traditional bank failure prediction literature in that it is a single-equation model, the primary goal of which is to predict bank failures; therefore, it does not formally distinguish between insolvency and failure. Thus, unlike the models in Thomson (1991) and Demirgüç-Kunt (1991), the one presented here does not allow for the study of bank closure policy. On the other hand, unlike the traditional failure prediction literature, this study includes proxy variables to control for the effects of Kane's four constraints on the probability of failure. Finally, the model is an extension of the previous failure prediction models in that it incorporates general measures of local economic conditions into the analysis.

The purpose of this study is to model bank failures of all sizes. This precludes the use of market data, which are available only for a limited number of large banking organizations. Therefore, I use proxy variables based on balance-sheet and income data from the call reports. These variables, defined in box 1, are drawn from the extensive literature on bank failures.

■ 3 For a review of this literature, see Demirgüç-Kunt (1989b).

■ 4 I consider a bank as failed if it is closed or requires FDIC assistance to remain open. For a discussion of the different failure resolution techniques available to the FDIC, see Caliguire and Thomson (1987).

BOX 1

Definitions of Proxy Variables

Dependent variable

DFAIL	Dummy variable: equals one for a failed bank, zero otherwise.

Regressors

NCAPTA	Book equity capital plus the reserve for loan and lease losses minus the sum of loans 90 days past due but still accruing and nonaccruing loans/total assets.
NCLNG	Net chargeoffs/total loans.
LOANHER	Loan portfolio Herfindahl index constructed from the following loan classifications: real estate loans, loans to depository institutions, loans to individuals, commercial and industrial loans, foreign loans, and agricultural loans.
LOANTA	Net loans and leases/total assets.
LIQ	Nondeposit liabilities/cash and investment securities.
OVRHDTA	Overhead/total assets.
ROA	Net income after taxes/total assets.
INSIDELN	Loans to insiders/total assets.
BRANCHU	Dummy variable: equals one if the state is a unit banking state, zero otherwise.
DBHC	Dummy variable: equals one if the bank is in a bank holding company, zero otherwise.
SIZE	Natural logarithm of total assets.
AVGDEP	Natural logarithm of average deposits per banking office.
BOUTDVH	Output Herfindahl index constructed using state-level gross domestic output by one-digit SIC codes.
UMPRTC	Unemployment rate in the county where the bank is headquartered.
CPINC	Percent change in state-level personal income.
BFAILR	Dun and Bradstreet's state-level small-business failure rate per 10,000 concerns.

The dependent variable, *DFAIL*, is the dummy variable for failure. The first eight regressors in the model are motivated by the early warning system literature. Early warning systems are statistical models for off-site monitoring of bank condition used by bank regulators to complement on-site examination. These models seek to determine the condition of a bank through the use of financial data.[5] The proxy variables used in the statistical monitoring models are motivated by the CAMEL rating categories, which regulators use during on-site examinations to determine a bank's condition. *NCAPTA*, the ratio of book equity capital less bad loans to total assets, is the proxy for capital adequacy (C̲AMEL). This variable is similar to Sinkey's (1977) net-capital-ratio variable, which is the ratio of primary capital less classified assets to total assets.[6] Both Sinkey and Whalen and Thomson (1988) show that similar proxy variables are better indicators of a bank's true condition than is a primary capital-to-assets ratio.

The next three early warning system variables are proxies for asset quality and portfolio risk (CA̲MEL). *NCLNG* measures net losses per dollar of loans and, hence, the credit quality of the loan portfolio. *LOANHER* is a measure of the diversification of the risky asset or loan portfolio and is therefore a measure of portfolio risk. *LOANTA* is the weight of risky assets in the total asset portfolio and, hence, a proxy for portfolio risk.

OVRHDTA and *INSIDELN* are proxies for management risk (CAM̅EL). *OVRHDTA* is a measure of operating efficiency, while *INSIDELN* is the proxy for another form of management risk: fraud or insider abuse. Graham and Horner (1988) find that for national banks that failed between 1979 and 1987, insider abuse was a significant factor, contributing to the failure of 35 percent of the closed institutions; material fraud was present in 11 percent of these failures. *ROA*, the return on assets, is the proxy for the earnings component of the CAMEL rating (CAME̲L), and *LIQ* is included to proxy for liquidity risk (CAMEL̲).

5 The purpose of early warning systems is to detect the deterioration of a depository institution's condition between scheduled examinations so that the FDIC can move that institution up in the on-site examination queue. For further information, see Korobow and Stuhr (1983), Korobow, Stuhr, and Martin (1977), Pettway and Sinkey (1980), Rose and Kolari (1985), Sinkey (1975, 1977, 1978), Sinkey and Walker (1975), Stuhr and Van Wicklen (1974), Wang and Sauerhaft (1989), and Whalen and Thomson (1988).

6 Classified assets is a measure of bad loans and other problem assets on a bank's confidential examination report; consequently, it is measured infrequently and is often unavailable to researchers.

In another study (Thomson [1991]), I show that *LOANTA, LIQ, OVRHDTA,* and *ROA* may also proxy for the non-solvency-related factors that contribute to the decision to close insolvent banks, providing additional justification for the inclusion of these variables in the failure prediction equation. I include the remainder of the variables listed in box 1 in the failure prediction equation either because the aforementioned study has shown them to be related to the closure decision (*BRANCHU, DBHC, SIZE, AVGDEP*), or because they serve as proxies for the economic conditions in the bank's home market (*BOUTDVH, UMPRTC, CPINC, BFAILR*).

BRANCHU is included in the regression to control for intrastate branching restrictions. Branching restrictions effectively limit both the opportunities for geographic diversification of a bank's portfolio and the FDIC's options for resolving an insolvency.

DBHC is a dummy variable for holding company affiliation, motivated by the source-of-strength doctrine. Source of strength is the regulatory philosophy, espoused by the Federal Reserve, that the parent holding company should exhaust its own resources in an attempt to make its banking subsidiaries solvent before asking the FDIC to intercede.

I include *SIZE,* the natural logarithm of total assets, in the failure prediction equation to control for the "too big to let fail doctrine" (TBLF). Bank regulators adopted TBLF in the 1980s as a result of the administrative difficulties, the implications for the FDIC insurance fund, and the political fallout associated with the failure of a large bank.

The average deposits per banking office, *AVGDEP,* is used as the proxy for franchise or charter value. Buser, Chen, and Kane (1981) argue that the FDIC uses charter values as a restraint on risk-taking by banks, and that bank closure policy is aimed at preserving charter value in order to minimize FDIC losses. Because the primary source of a bank's charter value is its access to low-cost insured deposits, the level of deposits per banking office should be positively correlated with the value of the banking franchise.

Finally, I include four measures of economic conditions in the bank's markets in order to incorporate the effects of local economic conditions on the bank's solvency: unemployment (*UMPRTC*), growth in personal income (*CPINC*), the business failure rate (*BFAILR*), and a measure of economic diversification (*BOUTDVH*). Unlike Gajewski (1989, 1990), who includes proxies for energy and agricultural shocks in his failure prediction models, I have included economic condition proxies that do not require knowledge of which economically important sectors will experience problems in the future.

II. The Data

Bank failures from July 1984 through June 1989 comprise the failed-bank sample and are taken from the FDIC's Annual Reports from 1984 through 1987 and from FDIC press releases for 1988 and 1989. Only FDIC-insured commercial banks in the United States (excluding territories and possessions) are included.

The nonfailed sample includes U.S. banks operating from June 1982 through June 1989 that filed complete call reports. I have drawn this sample randomly from the call reports and have checked the nonfailed sample to ensure that it is representative of the population of nonfailed banks. For instance, the majority of banks in the population are small; therefore, the nonfailed sample is drawn in a manner that ensures that small banks are adequately represented.

Data for the failed banks are drawn from the June and December call reports for 1982 through 1988 and are collected for up to nine semiannual reports prior to the date the bank was closed. I do not collect data for failed banks from call reports within six months of the failure date, because call reports are unavailable to regulators for up to 70 days after a report is issued. Furthermore, window dressing on the call reports of distressed banks just prior to their failure makes that data unreliable. In the cases where all or the majority of bank subsidiaries of a bank holding company are closed at once (for example, BancTexas Group, First City Bancorp of Houston, First Republic Bancorp of Dallas, and MCorp of Houston), the closed institutions are aggregated at the holding company level and treated as a single failure decision. I include a total of 1,736 banks in the nonfailed sample. The number of failed banks in the sample in each year appears in table 1.[7]

I obtain data on economic condition from several sources. State-level gross domestic output data are obtained from the Bureau of Economic Analysis for the years 1980 through

7 Call report data are screened for errors. I deleted from the sample those failed and nonfailed bank samples that were found to have missing or inconsistent loan data or negative values for expense items such as operating and income expense. Roughly 2 percent of the failed sample and 4 percent of the nonfailed sample were eliminated for these reasons. In addition, I eliminated banks in the nonfailed sample that were missing a June or December call report between 1982 and 1988.

TABLE 1

Number of Failed Banks in the Sample

Year	Number of banks
1984	78
1985	115
1986	133
1987	193
1988	174
1989[a]	77

a. 1989 failure number is for banks closed during the first six months of the year.
NOTE: Number of banks in the nonfailed sample in each year is 1,736.
SOURCES: Author's calculations and FDIC Annual Reports.

1986. County-level employment data are taken from the Bureau of Labor Statistics' *Employment and Earnings* for the years 1980 through 1986. State-level personal income data are from the Bureau of Economic Analysis' annual personal income files for the years 1981 through 1988, and business failure data are from Dun and Bradstreet's *Business Failure Record* for the years 1982 through 1988. Because all of the economic condition data are annual, I match the business failure and personal income data with the December call report data of the same year and the following June. The gross domestic output and employment data are matched with the December and June call report data in a similar manner, but with a two-year lag.[8]

III. Empirical Results

I estimate the model by logit regression using the logist regression procedure in SAS. I have chosen logit estimation rather than ordinary least squares (OLS) because of the undesirable properties of the OLS estimator when the dependent variable in the model is a binary (Amemiya [1981]). The unequal frequency of the failed and nonfailed samples suggests the use of logit rather than probit estimation because logit is not sensitive to the uneven sampling frequency problem (Maddala [1983]). The panel nature of the data allows two types of tests to be performed. First, I pool the data over time (using the June 1983 through December 1988 call reports) and assess the predictive accuracy of the model for up to 48 months before failure. Then, using the June call reports for 1983 through 1986, I ascertain the model's in-sample and out-of-sample accuracy.

Overall, the results indicate that up to 30 months before failure, solvency and liquidity are the most important predictors of failure. As the time to failure increases, however, asset quality, earnings, and management gain in importance as predictors of failure. The performance of the FDIC closure constraint proxies in table 2 demonstrates that the distinction between official failure and insolvency is significant and should be accounted for in studies of bank failures. Although the performance of the economic condition variables is mixed, their inclusion increases the predictive accuracy of the model.

Table 2 shows that the coefficient on *NCAPTA* is negative and significant for banks failing within 30 months of the call date and positive for banks failing within 30 to 48 months of the call date. However, the coefficient is only positive and significant for the 36- to 42-month subsample. The positive sign on *NCAPTA* for banks failing after 30 months is paradoxical, because it suggests that book solvency is positively related to failure. This, however, is not a new result (see Thomson [1991] and Seballos and Thomson [1990]). One possible explanation is that banks beginning to experience difficulties improve their capital positions cosmetically by selling assets on which they have capital gains and by deferring sales of assets on which they have capital losses. Another explanation, although not a mutually exclusive one, is that strong banks are more aggressive in recognizing and reserving against emerging problems in their loan portfolios than are weak banks.

The probability of failure is a negative function of asset quality, as the coefficient on *NCLNG* is negative and significant in all of the regressions except the six- to 12-month subsample. In addition, portfolio risk is positively related to the probability of failure, as evidenced by the positive and significant coefficient on *LOANTA* for all subsamples.

The positive and significant coefficients on *OVRHDTA* and *INSIDELN* for all subsamples indicate that management risk and insider

8 The output and employment data are matched with the sample having a two-year lag, because at the time this study was conducted I had access to these data only through the end of 1986; therefore, I could not match the employment and output data to the call data through 1988 without lagging them. Because a state's output mix is unlikely to change much in two years, this data-matching procedure should not affect the performance of *BOUTDVH*. However, while the decline of the financial sector is likely to follow a decline in the real sector, as measured by unemployment, the choice of a two-year lag is clearly ad hoc.

TABLE 2

Logit Regression Results from the Pooled Sample

	Months to failure after call report issued						
	6 to 12	12 to 18	18 to 24	24 to 30	30 to 36	36 to 42	42 to 48
φ_0	2.39	–0.16	–1.42	–3.27	–4.70	–4.52	–5.87
	(1.05)[a]	(0.95)	(0.93)	(0.95)[b]	(1.03)[b]	(1.12)[b]	(1.27)[b]
NCAPTA	–41.94	–30.71	–18.28	–11.08	1.06	4.72	7.21
	(1.73)[b]	(1.66)[b]	(1.56)[b]	(1.61)[b]	(1.73)	(1.68)[b]	(1.96)[b]
NCLNG	–5.08	–8.61	–14.63	–17.25	–23.92	–19.15	–19.10
	(3.72)	(4.13)[a]	(4.57)[b]	(5.07)[b]	(6.02)[b]	(6.52)[b]	(8.60)[b]
LOANHER	0.89	0.32	0.75	0.22	0.58	0.63	0.57
	(0.61)	(0.54)	(0.52)	(0.55)	(0.56)	(0.59)	(0.69)
LOANTA	6.84	8.25	8.96	9.36	9.47	8.60	8.34
	(0.60)[b]	(0.55)[b]	(0.53)[b]	(0.55)[b]	(0.58)[b]	(0.61)[b]	(0.70)[b]
OVRHDTA	193.37	192.71	234.77	261.39	229.73	242.42	282.60
	(28.04)[b]	(26.40)[b]	(25.55)[b]	(26.11)[b]	(27.79)[b]	(29.43)[b]	(34.27)[b]
INSIDELN	28.44	30.08	30.86	31.40	29.50	30.45	30.62
	(4.20)[b]	(3.67)[b]	(3.48)[b]	(3.52)[b]	(3.65)[b]	(3.85)[b]	(4.32)[b]
ROA	–43.87	–52.74	–62.87	–65.97	–82.64	–72.64	–78.57
	(5.69)[b]	(5.94)[b]	(6.55)[b]	(6.95)[b]	(8.11)[b]	(8.81)[a]	(10.81)[b]
LIQ	0.49	0.48	0.42	0.35	0.19	0.40	0.23
	(0.13)[b]	(0.13)[b]	(0.12)[b]	(0.12)[b]	(0.18)	(0.22)[c]	(0.36)
BRANCHU	0.05	0.14	0.10	–0.04	0.06	–0.03	0.30
	(0.14)	(0.13)	(0.12)	(0.12)	(0.13)	(0.14)	(0.15)[a]
DBHC	–0.66	–0.65	–0.66	–0.57	–0.50	–0.59	–0.59
	(0.13)[b]	(0.12)[b]	(0.11)[b]	(0.11)[b]	(0.12)[b]	(0.12)[b]	(0.15)[b]
SIZE	–0.79	–0.72	–0.68	–0.57	–0.45	–0.42	–0.18
	(0.11)[b]	(0.10)[b]	(0.09)[b]	(0.09)[b]	(0.10)[b]	(0.11)[b]	(0.12)
AVGDEP	0.23	0.37	0.43	0.49	0.45	0.45	0.29
	(0.14)[c]	(0.12)[b]	(0.11)[b]	(0.12)[b]	(0.12)[b]	(0.13)[b]	(0.15)[a]
BOUTDVH	–11.12	–12.88	–20.62	–20.70	–21.40	–22.08	–18.68
	(3.70)[b]	(3.27)[b]	(3.43)[b]	(3.53)[b]	(4.04)[b]	(4.38)[b]	(4.77)[b]
UMPRTC	–0.06	–0.04	–0.03	–0.04	–0.05	–0.07	–0.09
	(0.02)	(0.02)[a]	(0.02)[c]	(0.09)[a]	(0.02)[b]	(0.02)[b]	(0.03)[b]
CPINC	–12.19	–15.12	–13.95	–15.01	–14.19	–16.12	–19.33
	(2.58)[b]	(2.24)[b]	(2.07)[b]	(2.12)[b]	(2.23)[b]	(2.29)[b]	(2.66)[b]
BFAILR	–0.00	–0.00	–0.00	–0.00	–0.00	0.00	–0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.15)
χ^2[d]	3884.46[b]	2937.13[b]	2174.88[b]	1709.46[b]	1374.68[b]	1063.92[b]	854.93[b]
Type I	7.99	11.99	16.04	17.90	20.36	19.95	20.19
Type II	6.81	11.04	14.57	16.29	17.72	18.49	17.63
Class[e]	6.86	11.08	14.64	16.36	17.85	18.56	17.74
PPROB	0.04	0.04	0.05	0.05	0.05	0.05	0.05

a. Significant at the 5 percent level.
b. Significant at the 1 percent level.
c. Significant at the 10 percent level.
d. Model chi-square with 16 degrees of freedom.
e. Percentage of all banks misclassified.
NOTE: Dependent variable = *DFAIL*. Standard errors are in parentheses.
SOURCE: Author's calculations.

TABLE 3

Cross-Sectional Logit Regression Results

Call Date: Year Failed:	June 1984 1985	June 1985 1986	June 1986 1987
φ_0	0.41	0.54	1.38
	(2.85)	(2.88)	(2.34)
NCAPTA	–31.53	–29.90	–43.51
	(5.82)[a]	(4.75)[a]	(5.22)[a]
NCLNG	21.93	–1.21	7.20
	(20.01)	(13.72)	(8.77)
LOANHER	2.57	0.41	–0.92
	(1.39)[b]	(1.63)	(1.61)
LOANTA	10.13	9.49	7.37
	(1.74)[a]	(1.80)[a]	(1.52)[a]
OVRHDTA	301.45	242.80	489.16
	(86.83)[a]	(96.12)[c]	(86.14)[a]
INSIDELN	39.58	30.84	50.00
	(12.15)[a]	(10.75)[a]	(13.20)[a]
ROA	–49.91	–69.79	–0.41
	(27.18)[b]	(20.67)[a]	(17.66)
LIQ	2.71	1.53	0.95
	(1.22)[c]	(0.81)	(0.40)[c]
BRANCHU	0.11	–0.17	0.03
	(0.34)	(0.35)	(0.36)
DBHC	–0.79	–0.16	–0.99
	(0.34)[c]	(0.37)[a]	(0.33)[a]
SIZE	–1.13	–0.67	–0.91
	(0.32)[a]	(0.25)	(0.26)[a]
AVGDEP	0.58	0.42	0.50
	(0.38)	(0.32)	(0.32)
BOUTDVH	–10.22	–16.36	1.87
	(9.42)	(9.32)[b]	(8.24)
UMPRTC	–0.07	–0.03	–0.09
	(0.07)	(0.05)	(0.05)[c]
CPINC	–22.91	–30.19	–28.02
	(7.74)[a]	(8.81)[a]	(14.32)[b]
BFAILR	–0.00	–0.00	0.00
	(0.00)	(0.00)	(0.00)
χ^{2}[d]	402.59[a]	462.30[a]	667.31[a]
Type I	11.30	11.28	9.38
Type II	10.48	9.56	7.03
Class[e]	10.58	9.79	7.45
PPROB	0.13	0.15	0.22

a. Significant at the 1 percent level.
b. Significant at the 10 percent level.
c. Significant at the 5 percent level.
d. Model chi-square with 16 degrees of freedom.
e. Percentage of all banks misclassified.
NOTE: Dependent variable = *DFAIL*. Standard errors are in parentheses.
SOURCE: Author's calculations.

abuse are positively related to failure. In addition, the negative and significant coefficient on *ROA* in all subsamples and the positive and significant coefficient on *LIQ* for all regressions (except the 30- to 36-month and 42- to 48-month subsamples, for which the coefficient is positive and insignificant) indicate that the probability of failing is a negative function of earnings and liquidity.

With the exception of *BRANCHU* (all subsamples) and *SIZE* in the 42- to 48-month *SIZE* subsample, the coefficients on Thomson's (1991) closure constraint proxies are all significant, with the sign predicted by the author's call-option closure model in all the regressions.

The results for the economic condition variables are somewhat mixed. The coefficients on *BOUTDVH, UMPRTC,* and *CPINC* are negative and significant for all subperiods. In other words, the probability of failure is negatively related to state-level economic concentration *(BOUTDVH),* to county-level unemployment *(UMPRTC),* and to changes in state-level personal income *(CPINC).*

The negative sign on *CPINC* is consistent with its use as a proxy for differences between market and book solvency across regions. The significant negative relationship between the probability of failure and both *BOUTDVH* and *UMPRTC* is counterintuitive. If the condition of the banking industry were affected by the health of the economy, then I would expect the coefficients on both *BOUTDVH* and *UMPRTC* to be positive. *BOUTDVH* is a measure of economic diversity in the state where a bank does business. The more diversified a state's or region's economy, the more stable that economy should be and the lower *BOUTDVH* should be. It could be that *BOUTDVH* and *UMPRTC* are picking up the increased political constraints associated with the closing of banks in depressed regions like the Southwest. These political constraints increase as the number of insolvencies in a region grows. Finally, the coefficient on *BFAILR* is negative and insignificant for all subsamples.

Table 3 gives the results when the model is estimated using cross-sectional data from the June 1984, 1985, and 1986 call reports and from failures occurring in the subsequent calendar year. I use cross-sectional estimation for two reasons: 1) to test indirectly the pooling restriction imposed in the earlier tests and 2) to investigate the model's ability to predict failures outside the sample. To facilitate out-of-sample forecasting, I also split the nonfailed sample into two random samples of 868 banks. One is for use in in-sample forecasting, and the second is for use in out-of-sample forecasting. As seen in table 3, with the exception of the coefficients on *ROA* and *DBHC*, no significant difference seems to exist between the coefficients of each model across years. Therefore, the results reported in table 2 do not appear to be sensitive to the pooling restriction.

In-Sample Classification Accuracy

The second criterion for judging bank failure models is the classification accuracy of the model. In other words, how precise is the model in discriminating between failed and nonfailed banks within the sample, and how effective is it in discriminating between failed and nonfailed banks outside the sample?

For the pooled data, I perform only in-sample forecasting. Tables 2 and 3 report the overall classification accuracy of the three models, along with each model's type I and type II error. Type I error occurs when a failed bank is incorrectly classified as a nonfailed bank, and type II error occurs when a nonfailed bank is incorrectly classified as a failed bank. The overall classification error is the weighted sum of both types of errors. Typically, there is a trade-off between type I error and overall classification accuracy.

The logit model classifies a bank as failed if the predicted value of the dependent variable exceeds an exogenously set probability cutoff point (PPROB). The PPROB is set according to the prior probabilities of being in each group — typically, at 0.5. However, for studies such as this one, where closed banks are sampled at a higher rate than nonclosed banks, Maddala (1986) argues that the use of logit leads to a biased constant term that reduces the predictive power of the model. To correct for this, he suggests that one should assume that the prior probabilities are the sampling rates for the two groups. In addition, if type I error is seen to be more costly than type II error, a lower value for the PPROB is justified.

Overall, the model's in-sample classification accuracy is excellent (see table 2). Using the ratio of failed to nonfailed observations in the sample as the PPROB, I find that type I error ranges from 7.99 percent in the six- to 12-month subsample to 20.19 percent in the 42- to 48-month subsample. Overall classification error ranges from 6.86 percent in the six- to 12-month subsample to 18.56 percent in the 36- to 42-month subsample. As expected, type I errors and overall classification errors increase with time to failure.

Out-of-Sample Forecasting

One reason for studying bank failures is so that statistical models can be constructed to identify banks that may fail in the future. Such models are referred to as off-site monitoring or early warning systems in the literature and are used by bank regulators as a complement to on-site examinations. Out-of-sample forecasting not only yields information on the usefulness of the bank failure model as an examination tool, but also provides data on the stability of the failure equation over time.

For the out-of-sample forecasts, I use the estimated coefficients from the cross-sectional logit regressions, employing data from the June call reports of 1984 through 1986 and half of the nonfailed sample. The failed sample consists of all banks that failed in the year following the one from which the call report data were drawn. The coefficients for the model estimated over this sample appear in table 3. I use the second half of the nonfailed sample as the holdout sample for forecasting. I also construct three failed holdout samples using data from the June 1984 and June 1985 call reports. Only two holdout samples could be constructed for the June 1986 call date, because the failed-bank sample only runs through June 1989. The first failed holdout sample consists of banks failing in the second calendar year following the call report, and the second consists of banks failing in the third calendar year following the call report. The third holdout sample (unavailable for forecasting when the June 1986 call report is used) is comprised of banks failing in the fourth calendar year following the call report.

The results for this out-of-sample forecasting experiment appear in table 4. The PPROB cutoff point for classifying banks as failed or nonfailed is the ratio of failed to nonfailed banks from the in-sample regressions. Other cutoff points yield similar results. When PPROB = 0.132, the model misclassifies 10.19 percent of the banks in the holdout sample using 1986 failures. The type I error rate indicates that the model misclassifies nearly two-thirds of the failures, while roughly 2 percent of the nonfailed sample (type II error rate) is misclassified. Looking at the results for the 1987 and 1988 failure holdout samples, one can see that the type I errors and overall classification errors for all

TABLE 4

Out-of-Sample Forecasts

Date of call report	Failure date		
	1986	**1987**	**1988**
June 1984 (PPROB = 0.132)			
Type I	64.66	67.21	74.83
Type II	1.84	1.84	1.84
Class[a]	10.19	13.23	12.66
June 1985 (PPROB = 0.153)	**1987**	**1988**	**1989[b]**
Type I	63.73	71.93	75.34
Type II	1.73	1.73	1.73
Class[a]	13.01	13.28	7.44
June 1986 (PPROB = 0.221)	**1988**	**1989[b]**	
Type I	52.87	62.67	
Type II	3.23	3.23	
Class[a]	11.52	7.95	

a. Percentage of all banks misclassified.
b. 1989 sample of failed banks consists of banks closed during the first six months of the year.
NOTE: Forecasts employ the half of the nonfailed sample not used for the logit regressions in table 3.
SOURCE: Author's calculations.

TABLE 5

Additional Out-of-Sample Forecasts

Call date	Year failed	Type I	Type II	Class[a]
June 1985	1986	52.63	0.86	4.55
June 1986	1987	39.58	1.73	5.50
June 1987	1988	27.59	2.07	4.40
June 1988	1989[b]	22.08	1.67	2.54
In-sample forecast	—	12.48	10.21	10.91

a. Percentage of all banks misclassified.
b. 1989 sample of failed banks consists of banks closed during the first six months of the year.
NOTE: Out-of-sample forecasting is done with PPROB equal to 0.066 (the ratio of failed to nonfailed banks for the in-sample logit regressions) and using coefficients estimated from logit regressions on 1985 failures and the nonfailed sample from the June 1984 call report.
SOURCE: Author's calculations.

three models increase as we attempt to forecast further into the future. Note that the results for the June 1985 and June 1986 call reports are similar to those obtained using June 1984 data.

Given the high type I error rates, one might question the usefulness of the model for off-site monitoring of bank condition. However, the type I error rate could be lowered by decreasing PPROB enough so that the percentage of failed banks classified as nonfailed becomes acceptable. What is interesting from the standpoint of an early warning application is the low classification error and the low type II error. If one wanted to use this model to determine which banks should be examined next, low type II error would be an extremely important consideration, since the FDIC has limited examination resources.

In practice, the first out-of-sample experiment is of little use for designing early warning models because it requires the ability to identify failures in subsequent years in order to apply it. Therefore, I perform a second out-of-sample experiment that is able to mimic an early warning model in practice. Employing the June 1984 call report data, I estimate the three models using the entire nonfailed sample and the failures occurring in the next calendar year. I then use the coefficients to perform out-of-sample forecasting using 1) June call data for 1985 through 1988 on the nonfailed sample and 2) failures in the calendar year following the call report as the holdout samples. The PPROB is again set equal to the ratio of failed banks to nonfailed banks used in the in-sample logit regressions.

The results in table 5 show that using the 1984 version of the failure model, the out-of-sample classification error ranges from 5.50 percent in June 1986 to 2.54 percent in June 1988, and type I error ranges from 52.63 percent in June 1985 to 22.08 percent in June 1988. It is somewhat curious that the out-of-sample classification accuracy of all three models increases as we move further from the call date of the in-sample experiment. Again, note that the type I error for the out-of-sample regressions could be lowered at the expense of the type II error (and the overall classification error) by decreasing PPROB. The performance of all the models in the second out-of-sample forecasting experiment suggests that they could be used as part of an early warning system of failure.

IV. Conclusion

This study shows that the probability that a bank will fail is a function of variables related to its solvency, including capital adequacy, asset quality, management quality, earnings performance, and the relative liquidity of the portfolio. In fact, the CAMEL-motivated proxy variables for bank condition demonstrate that the majority of these factors are significantly related to the probability of failure as much as four years before a bank fails.

Overall, the model demonstrates good classification accuracy in both the in-sample and out-of-sample tests. For the in-sample tests, it is able to classify correctly more than 93 percent of the banks in the six- to 12-month subsamples and more than 82 percent of the banks in the 42- to 48-month subsamples. In addition, the model correctly classifies more than 94 percent of those banks that fail between six and 12 months of the call date and almost 80 percent of those that fail between 42 and 48 months of the call date. Out-of-sample classification accuracy is also excellent, indicating that the model could be modified for use as an early warning model of bank failure.

Economic conditions in the markets where a bank operates also appear to affect the probability of bank failure as much as four years before the failure date. However, given that regional economic risk is diversifiable, the sensitivity of the banking system to regional economic conditions suggests that policymakers should revise the laws and regulations that limit banks' ability to diversify their portfolios geographically (especially in light of the fact that the national economy was relatively strong during the years covered in this study).

Finally, the performance of the closure-constraint proxy variables indicates that the probability of failure is not simply the probability that a bank will become insolvent, but that it will be closed when it becomes insolvent. In other words, the results show that the distinction between official failure and economic insolvency is an important one, suggesting the need for further research on the determinants of the incentive systems faced by bank regulators (see Kane [1986, 1989]).

References

Amemiya, T. "Qualitative Response Models: A Survey," *Journal of Economic Literature,* vol. 19 (December 1981), pp. 1483–1536.

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

Bovenzi, John F., and Arthur J. Murton. "Resolution Costs of Bank Failures," *FDIC Banking Review,* vol. 1, no. 1 (Fall 1988), pp. 1–13.

Buser, Stephen A., Andrew H. Chen, and Edward J. Kane. "Federal Deposit Insurance, Regulatory Policy, and Optimal Bank Capital," *Journal of Finance,* vol. 36, no. 1 (March 1981), pp. 51–60.

Caliguire, Daria B., and James B. Thomson. "FDIC Policies for Dealing with Failed and Troubled Institutions," Federal Reserve Bank of Cleveland, *Economic Commentary,* October 1, 1987.

Demirgüç-Kunt, Aslı. "Modeling Large Commercial-Bank Failures: A Simultaneous-Equation Analysis," Federal Reserve Bank of Cleveland, Working Paper 8905, May 1989a.

______. "Deposit-Institution Failures: A Review of Empirical Literature," Federal Reserve Bank of Cleveland, *Economic Review,* 4th Quarter 1989b, pp. 2–18.

______. "On the Valuation of Deposit Institutions," Federal Reserve Bank of Cleveland, Working Paper 9104, March 1991.

______. "Principal–Agent Problems in Commercial-Bank Failure Decisions," Federal Reserve Bank of Cleveland, Working Paper (forthcoming).

Gajewski, Gregory R. "Assessing the Risk of Bank Failure," in *Proceedings from a Conference on Bank Structure and Competition,* Federal Reserve Bank of Chicago, May 1989, pp. 432–56.

______ . "Modeling Bank Closures in the 1980's: The Roles of Regulatory Behavior, Farm Lending, and the Local Economy," in George G. Kaufman, ed., *Research in Financial Services: Private and Public Policy.* Greenwich, Conn.: JAI Press, Inc., 1990, pp. 41–84.

Graham, Fred C., and James E. Horner. "Bank Failure: An Evaluation of the Factors Contributing to the Failure of National Banks," in *Proceedings from a Conference on Bank Structure and Competition,* Federal Reserve Bank of Chicago, May 1988, pp. 405–35.

Kane, Edward J. "Appearance and Reality in Deposit Insurance: The Case for Reform," *Journal of Banking and Finance,* vol. 10 (June 1986), pp. 175–88.

______ . *The S&L Insurance Mess: How Did It Happen?* Washington, D.C.: The Urban Institute, 1989.

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

______ , ______ , and Daniel Martin. "A Nationwide Test of Early Warning Research in Banking," Federal Reserve Bank of New York, *Quarterly Review,* Autumn 1977, pp. 37–52.

Lane, William R., Stephen W. Looney, and James W. Wansley. "An Application of the Cox Proportional Hazards Model to Bank Failure," *Journal of Banking and Finance,* vol. 10, no. 4 (December 1986), pp. 511–31.

______ , ______ , and ______ . "An Examination of Bank Failure Misclassification Using the Cox Model," in *Proceedings from a Conference on Bank Structure and Competition,* Federal Reserve Bank of Chicago, May 1987, pp. 214–29.

Maddala, G.S. *Limited-Dependent and Qualitative Variables in Econometrics.* New York: Cambridge University Press, 1983.

______ . "Econometric Issues in the Empirical Analysis of Thrift Institutions' Insolvency and Failure," Federal Home Loan Bank Board, Invited Research Working Paper 56, October 1986.

Pettway, Richard H., and Joseph F. Sinkey, Jr. "Establishing On-Site Bank Examination Priorities: An Early Warning System Using Accounting and Market Information," *Journal of Finance,* vol. 35 (March 1980), pp. 137–50.

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

Seballos, Lynn D., and James B. Thomson. "Underlying Causes of Commercial Bank Failures in the 1980s," Federal Reserve Bank of Cleveland, *Economic Commentary,* September 1, 1990.

Sinkey, Joseph F., Jr. "A Multivariate Statistical Analysis of the Characteristics of Problem Banks," *Journal of Finance,* vol. 30 (March 1975), pp. 21–36.

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

______ . "Identifying 'Problem' Banks: How Do the Banking Authorities Measure a Bank's Risk Exposure?" *Journal of Money, Credit and Banking,* vol. 10 (May 1978), pp. 184–93.

______ , and David A. Walker. "Problem Banks: Definition, Importance and Identification," *Journal of Bank Research,* vol. 5, no. 4 (Winter 1975), pp. 209–17.

Stuhr, David P., and Robert Van Wicklen. "Rating the Financial Condition of Banks: A Statistical Approach to Aid Bank Supervision," Federal Reserve Bank of New York, *Monthly Review,* September 1974, pp. 233–38.

Thomson, James B. "An Analysis of Bank Failures: 1984 to 1989," Federal Reserve Bank of Cleveland, Working Paper 8916, December 1989.

______. "Modeling the Bank Regulator's Closure Option: A Two-Step Logit Regression Approach," unpublished manuscript, January 1991.

Wang, George H. K., and Daniel Sauerhaft. "Examination Ratings and the Identification of Problem/Non-Problem Thrift Institutions," *Journal of Financial Services Research,* vol. 2 (October 1989), pp. 319–42.

West, Robert Craig. "A Factor-Analytic Approach to Bank Condition," *Journal of Banking and Finance,* vol. 9 (June 1985), pp. 253–66.

Whalen, Gary. "A Proportional Hazards Model of Bank Failure: An Examination of Its Usefulness as an Early Warning Tool," Federal Reserve Bank of Cleveland, *Economic Review,* 1st Quarter 1991, pp. 21–31.

______**, and James B. Thomson.** "Using Financial Data to Identify Changes in Bank Condition," Federal Reserve Bank of Cleveland, *Economic Review,* 2nd Quarter 1988, pp. 17–26.