FDIC Banking Review

2003 Volume 15, No. 4

Options for Pricing Federal Deposit Insurance

*Eric P. Bloecher, Gary A. Seale, and Robert D. Vilim**

Editors Note:

The following article discusses some deposit insurance pricing options that are under consideration by the FDIC. The specific pricing examples are presented only to illustrate the general types of options being considered and should not be regarded as a comprehensive set. This article is intended to highlight the practical trade-offs posed by the choice among different types of pricing systems.

In April 2001, the FDIC released a document entitled "Keeping the Promise: Recommendations for Deposit Insurance Reform" (the recommendations paper), which laid out the Corporation's recommendations for merging the insurance funds, eliminating the designated reserve ratio as the trigger for charging premiums, considering rebates if the merged fund grows too rapidly, and indexing insurance coverage. The paper also recommended charging regular, risk-based insurance premiums to all banks,[1] and it included some examples of how the FDIC might enhance the current nine-cell premium matrix (see table 1 in the next section) to better price for risk.

Since the FDIC released the recommendations paper, our work has focused on further exploring the options for pricing deposit insurance. Generally, we have been reviewing three primary methodologies: expanded use of supervisory ratings, use of statistical models, and a combination of the two. Choosing a system for deposit insurance pricing involves trade-offs among a number of desirable attributes. We summarize the options being explored and discuss the trade-offs without offering a judgment as to which attributes of deposit insurance pricing are most desirable from a policy standpoint.

* All the authors are with the Division of Insurance and Research at the Federal Deposit Insurance Corporation. Eric P. Bloecher and Robert D. Vilim are senior financial analysts, and Gary A. Seale is a senior financial economist. The authors would like to thank John O'Keefe and Richard Jones for helpful comments, and Christine Brickman for research assistance.

[1] Throughout this article, the term *banks* refers to all FDIC-insured institutions.

After providing some historical background on FDIC premiums, we lay out the desirable attributes, or general requirements, of deposit insurance pricing; present the pricing options we have been considering, along with the relative merits of each option; and then describe pricing for two categories of banks that could be priced separately: new banks and large banks. The final section concludes with a brief summary of the trade-offs that need to be evaluated before a new deposit insurance pricing system is selected.

Historical Background

For most of the FDIC's history, deposit insurance coverage was funded by a premium system under which all insured institutions were charged an identical flat rate for deposit insurance. The rate was set by the Banking Act of 1935 as 1/12 of 1 percent of total domestic deposits.[2] Thus, deposit insurance premiums did not vary with the level of risk that an institution posed to the insurance fund.

After passage of the Banking Act, the banking industry stabilized quickly, and bank failures remained low through the 1940s. The rapid increase in lending after the war was not accompanied by the high loan losses that many had anticipated; instead, the FDIC was faced with the possibility that the insurance fund could grow unchecked. To address this issue, Congress passed the Federal Deposit Insurance Act of 1950, which provided for assessment credits to be distributed to banks in years when the FDIC's assessment income exceeded its losses and expenses. The credits were distributed on a pro rata basis, with the FDIC retaining up to 40 percent of the Corporation's net assessment income and banks receiving up to 60 percent. The system of credits was a way to control the growth of the insurance fund by allowing premium income to be reduced in periods with low failure rates, while the FDIC retained the ability to make full use of premiums during periods of higher failure rates.

Although many observers recognized from the beginning that the original pricing system had weaknesses, the full implications of flat-rate insurance assessments did not attract significant attention until the bank and savings and loan insurance funds experienced record losses in the late 1980s. Two main problems were identified. First, a flat-rate system provided an inducement for a bank or thrift to undertake higher-risk business strategies to maximize profits. These strategies could be pursued without the banks incurring additional insurance expense; failure costs generated by increased risk taking were instead passed on to the insurer (and perhaps the taxpayer). Second, in a flat-rate premium system, sound and well-managed institutions were subsidizing high-risk, poorly managed institutions: low-risk banks paid more for insurance than they should, whereas risky banks paid less. The subsidy funded by low-risk banks represented an economic burden that caused them to operate at a competitive disadvantage. These two problems pointed to the conclusion that a more equitable and economically supportable deposit insurance pricing system would require high-risk institutions to pay more than low-risk institutions.

The Federal Deposit Insurance Corporation Improvement Act of 1991 (FDICIA) required that a risk-based premium system be implemented by January 1, 1994. The FDIC implemented a risk-based system on January 1, 1993, a year early. Separate but identical assessment rate schedules were adopted for the Bank Insurance Fund (BIF) and the Savings Association Insurance Fund (SAIF). Institutions were assigned to one of nine risk categories by the use of capital ratios and other relevant information, mainly supervisory ratings. Originally, assessment rates ranged from 23 cents per $100 of assessable deposits for the lowest-risk institutions to 31 cents per $100 of assessable deposits for the highest-risk institutions. When the funds were recapitalized, premiums were lowered. The Deposit Insurance Funds

[2] This rate was calculated to be the annual assessment rate that would have been required to cover actual losses on deposits in banks that failed between 1865 and 1934, excluding "crisis" years when losses were unusually high.

Act of 1996 prohibits the FDIC from charging premiums to institutions that are well capitalized and highly rated by supervisors as long as the insurance fund is above 1.25 percent of insured deposits. Table 1 presents the current nine-cell matrix for the combined fund—that is, a hypothetical fund in which the BIF and SAIF are merged—and gives the number and percentage of banks in each cell as of year-end 2002. As the table indicates, over 90 percent of institutions are in the 1A category. Currently, these institutions are not assessed for deposit insurance.

Table 1

Matrix Distribution, Risk-Related Premium System
(Bank Insurance Fund and Savings Association Insurance Fund Combined, Year-End 2002)

	Supervisory Subgroup[b]		
Capital Subgroup[a]	**A (CAMELS 1 or 2)**	**B (CAMELS 3)**	**C (CAMELS 4 or 5)**
1–Well Capitalized	8,583 91.7%	523 5.6%	115 1.2%
2–Adequately Capitalized	113 1.2%	17 0.2%	14 0.1%
3–Undercapitalized	1 0.0%	0 0.0%	6 0.1%

Note: The figures in the cells refer to the number and percentage of all FDIC-insured institutions.

[a] Assignments to capital subgroups are made in accordance with section 327.4(a)(1) of the FDIC's Rules and Regulations. "Well capitalized" means a total risk-based capital ratio that equals or exceeds 10 percent, a Tier-1 risk-based capital ratio that equals or exceeds 6 percent, and a Tier-1 leverage capital ratio that equals or exceeds 5 percent. "Adequately capitalized" means not well capitalized and a total risk-based capital ratio that equals or exceeds 8 percent, a Tier-1 risk-based capital ratio that equals or exceeds 4 percent, and a Tier-1 leverage capital ratio that equals or exceeds 4 percent. "Undercapitalized" means neither well capitalized nor adequately capitalized.

[b] Assignments to supervisory subgroups are made in accordance with section 327.4(a)(2) of the FDIC's Rules and Regulations. Subgroup A consists of financially sound institutions that have only a few minor weaknesses; this subgroup generally corresponds to the primary federal regulator's composite CAMELS rating of 1 or 2. Subgroup B consists of institutions with demonstrable weaknesses that, if not corrected, could lead to a significant deterioration of the institution and an increased risk of loss to the relevant insurance fund; this subgroup generally corresponds to the primary federal regulator's composite CAMELS rating of 3. Subgroup C consists of institutions that pose a substantial probability of loss to the relevant insurance fund unless effective corrective action is taken. This subgroup generally corresponds to the primary federal regulator's composite CAMELS rating of 4 or 5.

Key Attributes of a Deposit Insurance Pricing Structure

Ideally, any pricing system adopted by the FDIC would possess some combination of five attributes: accuracy, simplicity, flexibility, appropriate incentives, and fairness.

Accuracy

Perhaps the most important consideration for any proposed pricing system is that the criteria used to rank or categorize banks accurately reflect the relative risk that institutions pose to the insurance fund. Accuracy is generally measured against the insurable event, which in this case is bank failure.[3] Banks that are in higher-premium categories should have a more frequent occurrence of failure than banks in lower-premium categories.

Additionally, for any pricing methodology that relies extensively on data provided by banks or other outside parties, the integrity of the data must be adequate. Reported data must be timely, accurate, and verifiable. They must be available to regulators early enough in the assessment cycle to allow for premiums to be calculated.

Simplicity

The methodology selected should be available to the public, insured banks, and other outside parties, and members of all three groups should find it comprehensible. Moreover, bankers should be able to compute their risk categories or ratings without undue difficulty—preferably, early in the assessment cycle. For some pricing systems, the FDIC may need to provide software or some other form of technical assistance to help bankers perform the calculations.

[3] For certain groups of banks or within certain time periods, data on failures are often insufficient to allow meaningful statistical comparisons. As a result, to compare the pricing methodologies considered here, we also use historical data on the frequency of examination rating downgrades.

Flexibility

The factors that are most predictive of bank failure can change over time. Moreover, we expect that the FDIC's ability to measure risk exposure will improve over time. Consequently, it is important that any pricing system allow for periodic changes in the risk-assessment criteria. Allowing for periodic changes will enable the FDIC to continually evaluate which factors are relevant at any particular time. Updates should be infrequent enough to allow banks a measure of stability for planning purposes, yet frequent enough to ensure that all the criteria remain relevant.

Appropriate Incentives

The measures of bank risk included in the pricing structure should provide incentives for bank management to act responsibly. While some measures of risk may perform well in statistical tests, their inclusion in a pricing system may not be appropriate because of the perverse incentives they create for sound bank management. For example, a measure that penalized banks for increasing levels of charge-offs might create an incentive for managers to avoid charging off loans simply to reduce their insurance premiums.

Fairness

Closely associated with the idea that the classification is to be correlated with risk is the idea that banks with similar characteristics should be treated in a like manner. Institutions with similar risk structures should pay approximately equal premiums.

Pricing Options for Well-Capitalized and Highly-Rated Institutions

A primary objective of deposit insurance pricing reform is to better differentiate among the best-rated institutions on the basis of risk, thereby reducing subsidies paid by low-risk institutions to riskier ones and moderating incentives for increased risk taking.[4] Pricing that incorporates greater sensitivity to risk would achieve the goal of making the deposit insurance system more equitable and economically efficient.[5]

The options currently being considered for banks (other than large banking organizations) include expanded use of supervisory ratings, use of statistical models (both in a continuous and discrete format), use of a combination of statistical models and supervisory ratings, and a scorecard that uses expert judgment in conjunction with a statistical model.

Expanded Use of Supervisory Ratings

A simple method of providing further risk differentiation within the best insurance category is to make expanded use of the CAMELS ratings.[6] This expansion could involve either the use of composite ratings alone or the use of composite and component ratings combined. If composite ratings alone are used, composite 1-rated institutions would pay a lower premium than composite 2-rated institutions. Table 2 presents examples of how the FDIC might use both the composite rating alone and the composite rating combined with the component ratings to subdivide the 1A

[4] Because of the statutory prohibition noted above, currently subsidies are paid only when the insurance fund is less than 1.25 percent of insured deposits. For purposes of this article, we concentrate on banks within the best insurance category (the 1A category). Institutions that are not well capitalized and highly rated are generally subject to a higher level of supervisory review and, in some cases, may be operating under specific enforcement actions.

[5] This article is concerned primarily with differentiating banks according to the risk they pose to the insurance funds, not with determining the absolute amounts that individual banks should pay for insurance (that is, not with determining the "break-even" or "actuarially fair" amounts). Actuarial pricing is the goal of most private insurers and as a general approach has much to recommend it, but adopting a strict actuarial framework would be impractical for the FDIC, mainly because if the FDIC were to charge the highest-risk institutions such a premium, the premium would be high enough to threaten these banks with failure.

[6] The CAMELS rating is assigned by a bank's primary regulator. The acronym stands for **C**apital, **A**ssets, **M**anagement, **E**arnings, **L**iquidity, and **S**ensitivity to market risk. A rating from 1 (the best) to 5 (the worst) is assigned for each of these component elements, and an overall composite rating based on the component ratings is then assigned to the bank.

insurance category. In option 1 of the table (composite ratings only), banks rated a composite 1 are placed in the lower-premium category (1A1) and banks rated composite 2 are placed in the higher-premium category (1A2). As of year-end 2002, 40.4 percent of institutions would have fallen in the 1A1 category and 59.6 percent of institutions would have fallen in the 1A2 category. In option 2 (composite and component ratings combined) the institutions rated composite 1 are placed in the 1A1 category as in option 1. However, composite 2-rated banks would be divided into two groups based on their component ratings: most would categorized into the 1A2 category, while those banks having weaker component ratings would be placed into the 1A3 category. Banks in the 1A3 category would pay the highest premium rates among the 1A banks.

Table 2

Options for Pricing Well-Capitalized and Highly-Rated Institutions Using CAMELS Ratings (Year-End 2002)

	Subcategories 1A1	1A2	1A3
Option 1: Using Composite Ratings			
Composite 1 Rated	3,501 40.4%		
Composite 2 Rated		5,169 59.6%	
Option 2: Using Composite and Component Ratings			
Composite 1 Rated	3,501 40.4%		
Composite 2 Rated *and* Sum of Components <= 12 *and* No More Than One Component Rated 3 or Worse		4,271 49.3%	
Composite 2 Rated *and* Sum of Components > 12 *or* Two or More Components Rated 3 or Worse			898 10.4%

Note: The table shows two options for subdividing the 1A insurance category using supervisory ratings. The 1A1 subgroup represents the least risk and the 1A3 subgroup represents the greatest risk. The figures in the cells refer to the number and percentage of 1A institutions.

Statistical Models

Pricing methods that rely on statistical models have been developed to provide options that incorporate objective financial data reported by banks. The two statistical models under consideration use reported financial data to rank banks in the 1A category. One is a failure-prediction model, and the other is a CAMELS downgrade-prediction model.

Failure-Prediction Model

The failure-prediction model is a statistical model that relates historical Call Report ratios to bank failures to determine an estimated failure probability for each bank.[7] This failure probability can be used to rank banks for pricing. Table 3 illustrates an example of a failure-prediction calculation for a hypothetical bank. Column A shows the coefficients produced by the model. These coefficients are the same for all banks and represent the relative weight placed on each ratio for determining a probability of failure. The hypothetical bank's financial ratios, which can be obtained from the Call Report, are in column B. These ratios are multiplied by the corresponding coefficients in column A to obtain the values in column C. The sum of these values produces a raw score, which is then transformed to obtain the estimated failure probability.[8] For the hypothetical bank, this probability is 0.39 percent.

The estimated probability of failure for each bank ranges between 0 and 100 percent. This value represents the likelihood of a bank's failing over a five-year period. Under a continuous pricing format, in which each institution receives an individual score, banks could be ranked according to their estimated failure probabilities and assessed according to their ranking.

[7] The model is a logistic model of the general form $\rho\{1|X,\beta\} = e^Z/(1 + e^Z)$ where $Z = \alpha + \Sigma\beta_i x_i$, the number 1 represents bank failure within a specified period, and x_i represents the *i* th financial-ratio variable.

[8] The transformation follows the formula in the above footnote.

Table 4 shows how failure probabilities from the statistical failure model would have been distributed among all well-capitalized and highly-rated FDIC-insured banks for selected years between 1985 and 2002. Most banks would have had a low estimated probability of failure, especially after 1995. Thus, it is reasonable to assume that during periods of relative stability, most banks would pay an amount close to the average premium for that category.

Table 3

Failure-Prediction Model, Hypothetical Bank

Scoring Factor	Coefficient (Weight) (A)	Financial Ratio (B)	Score[a] (C)
Intercept	–3.91	N/A	–3.91
Nonaccrual Loans / Total Assets	35.47	.002	0.07
Loans Past Due 90+ Days / Total Assets	37.10	.010	0.37
ORE / Total Assets	30.46	.015	0.45
Loans Past Due 30–89 days / Total Assets	30.45	.005	0.15
Pretax Net Operating Income / Average Assets	–15.17	.030	–0.45
Noncore Funding / Total Assets	5.20	.120	0.62
Equity & Reserves / Total Assets	–21.69	.130	–2.82
Total Score			**–5.52**

Note: This table demonstrates how the results of the failure-prediction model can be used to create an individual expected-failure probability for each institution.

[a] The raw score is the product of columns A and B. Via the formula in note 7, the total score produces the expected probability of failure (Pr(default)) through the transformation.

$$\text{Pr(default)} = \frac{e^{-5.52}}{(1 + e^{-5.52})} = 0.39\%$$

Figures 1 and 2 show how the failure-prediction model would have performed historically in identifying both CAMELS downgrades (figure 1) and failures (figure 2). In figure 1, we used Call Report data at each year-end to rank banks according to their expected failure probabilities; we then divided the rank listing into three numerically equal groups and, for each group, calculated the percentage of banks that were actually downgraded from a CAMELS 1 or 2 to a CAMELS 3, 4, or 5 over the subsequent five-year period. As the figure shows, the group with the highest expected failure rate consistently has the highest percentage of banks downgraded. Likewise, banks in the middle group of the three consistently have a higher percentage of downgrades than banks in the group with the lowest expected rate of failure.

Figure 2 shows the percentage of banks in each of the three groups that actually failed over the subsequent five-year period. This figure, too, shows a consistently higher failure rate for the group of banks having the highest expected failure rate. The distinction is not as clear for the middle and lowest thirds, however, primarily because of the low overall number of failures in these groups, especially after 1992. (This problem of the low overall number of actual failures distorting the percentages after 1992 is common to all the pricing systems evaluated here.)

Table 4

Continuous Pricing Distribution of Well-Capitalized and Highly-Rated Institutions Based on Projected Failure Probabilities

	Projected Range of Failure Probability									
	<= 0.5	0.5–1.0	1.0–1.5	1.5–2.0	2.0–2.5	2.5–3.0	3.0–3.5	3.5–4.0	4.0–4.5	> 4.5
2002	42.3%[a]	28.1%	12.7%	5.8%	3.0%	2.0%	1.3%	0.9%	0.7%	3.2%
2000	42.5	27.4	12.6	5.6	3.5	2.3	1.3	0.9	0.7	3.3
1995	55.5	27.2	8.5	3.0	1.8	1.0	0.7	0.4	0.3	1.6
1990	36.5	29.7	12.8	6.6	3.9	2.2	1.7	1.1	0.9	4.7
1985	28.6	27.4	14.1	7.4	5.2	3.0	2.2	1.7	1.4	8.9

Note: Data are as of year-end.

[a] The percentages are those of 1A institutions in each of the failure-probability ranges. These percentages are based on the model shown in table 3.

CAMELS Downgrade-Prediction Model

The **S**tatistical **C**AMELS **O**ff-site **R**ating (SCOR) model is similar to the failure-prediction model but was designed specifically to estimate the likelihood that a bank will receive a CAMELS downgrade over the next year. The FDIC currently uses this model for the off-site risk monitoring of banks. The model produces an expected CAMELS rating for the bank, which is expressed as a number between 1.00 (the best) and 5.00 (the worst). The SCOR rating could be used to rank banks by risk for pricing purposes. In historical tests of downgrades and failures, SCOR performs much like the failure-prediction model, producing results very similar to those shown in figures 1 and 2.

Continuous versus Discrete

Separate from the choice of whether supervisory ratings or statistical models should be used to rank institutions is the question of whether a discrete or continuous format should be used. The failure-prediction model and the SCOR model produce a continuous ranking. (After each bank receives an individual score based on the results of the models, premium amounts are established on the basis of the relative ranking of each bank.) However, it is possible to create a discrete pricing structure by superimposing a fixed number of categories on the results of the models; for example, to create figures 1 and 2 we arbitrarily divided the banks into three groups with an equal number of institutions in each group. But the groups do not necessarily have to be of equal size. Rather, groups could be established that minimized the difference in expected failure probabilities between the best and worst banks in each group. Doing this is desirable, since grouping makes it inevitable that some banks will pay a somewhat higher premium than their expected failure probabilities will warrant, while others will pay a somewhat lower premium.

Discrete formats may offer greater simplicity than continuous formats, but they also create the potential that small changes in a measured variable could produce large changes in the deposit insurance premium ("cliff effects"). The existing nine-cell matrix is an example of a discrete format with cliff effects; however, it is based on well-established and generally accepted thresholds.

Figure 1

Relationship between Probability of Failure and Actual CAMELS Downgrades within Five Years of Failure-Probability Calculation

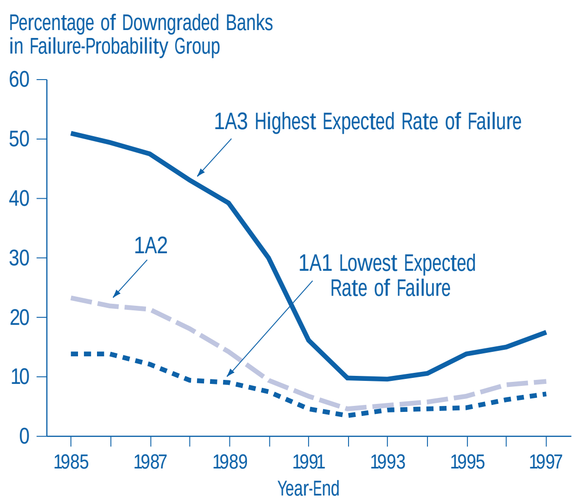
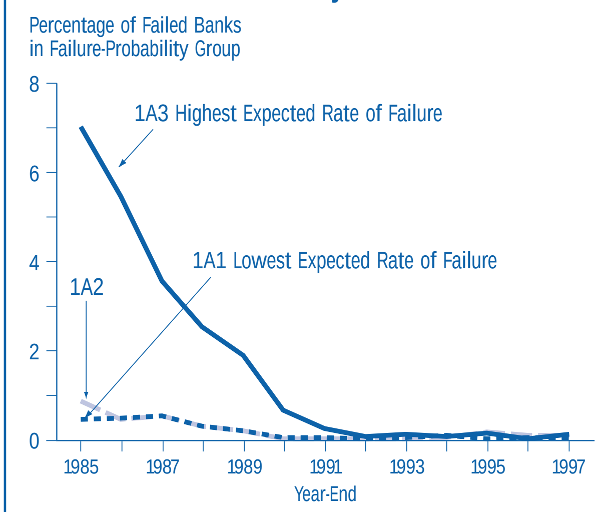


Figure 2

Relationship between Probability of Failure and Actual Failures within Five Years of Failure-Probability Calculation



For example, there is widespread understanding of the difference between a CAMELS 2-rated institution and a CAMELS 3-rated institution, and there is even different supervisory treatment for the two. Additionally, the capital thresholds are the same as the thresholds established by regulation for purposes of Prompt Corrective Action.

The problem of creating cliff effects can be mitigated by the use of a more graduated pricing structure. Any of the statistical methods could achieve this result. However, the greater the number of categories created, the more stringent the requirements that are placed on the specified system. If too many categories are created, the distinctions between them become less clear, and fairness becomes an issue. This is also the case with continuous pricing systems, when the number of categories is essentially equal to the number of banks. The methodologies described above could be used to create as many or as few pricing categories as required to achieve an acceptable trade-off, but it is important that expected failure rates be progressively higher for successively higher risk-pricing categories.

Combination of Supervisory Ratings and Statistical Models

Although using statistical models to price deposit insurance premiums might be appealing because of their reliance on objective financial data reported by banks, strictly applying a statistical model would inevitably result in some CAMELS 1-rated institutions paying more than some CAMELS 2-rated institutions. Since banks rated a composite 2 fail more frequently than banks rated a composite 1, it would seem logical to make the case that the 2-rated institutions in general should pay a higher premium for deposit insurance.

An alternative that addresses this concern is to combine supervisory ratings with one of the statistical models. A combined approach would preserve the CAMELS rating by initially classifying banks according to whether they were ranked CAMELS 1 or 2 and would then use the statistical models to create subcategories.

There are a number of potential possibilities for combining supervisory ratings and statistical models. Table 5 illustrates one way in which a pricing system might operate using CAMELS ratings and SCOR. In the table, we divide the 1A insurance category as of year-end 2002 into three subcategories. All CAMELS 1-rated banks are placed in the 1A1 category. CAMELS 2-rated banks with a SCOR value of less than 1.25 also are placed in the 1A1 category. The remaining CAMELS 2-rated banks are classified as either 1A2 (SCOR rating anywhere from 1.25 to 1.75) or 1A3 (SCOR rating higher than 1.75). The distributions shown in Table 5 would vary depending on the threshold values chosen for the SCOR ratings.

Combined approaches tend to perform much like the statistical methodologies in identifying downgrades and failures. Banks that are in higher-premium groups are more likely to be downgraded or to fail than banks in lower-premium groups.

Table 5

Option for Pricing Well-Capitalized and Highly-Rated Institutions Using CAMELS and SCOR Ratings (Year-End 2002)

	Subcategories		
	1A1	**1A2**	**1A3**
All Composite 1 Rated *and* Composite 2 Rated with SCOR Rating < 1.25	3,618 41.7%		
Composite 2 Rated with SCOR Rating >= 1.25 *and* <= 1.75		2,767 31.9%	
Composite 2 Rated with SCOR Rating > 1.75			2,285 26.4%

Note: SCOR values are calculated from December 31, 2002, Call Report data, and these values are combined with December 2002 exam ratings. The 1A1 subgroup represents the least risk and the 1A3 subgroup represents the greatest risk. The figures in the cells refer to the number and percentage of 1A institutions.

Scorecard

The scorecard uses an expert system to develop gradations of risk for each variable in the failure-prediction model, thus lessening the cliff effects. The original example of a scorecard appeared in the recommendations paper in April 2001. Since then, we have held numerous meetings with other regulators, industry groups, and academics to solicit ideas on the scorecard. The comments received from these groups led us to make adjustments to correct for criteria that unduly penalized a particular class of banks. Other changes were designed to improve the estimation techniques that had been used to create the original scorecard. The most significant change is that the failure-prediction model was reestimated for banks in the 1A insurance category only, rather than for the entire industry.

Table 6 shows the most recent version of the scorecard. In this example, the 1A category of the current pricing matrix (table 1) is divided into three subcategories. The scoring framework allows banks to be classified as 1A1 (least risk), 1A2, or 1A3 (most risk). The noncore-funding adjustment factor at the bottom of the table is included to address the unique funding strategies used by large banks and is discussed more fully below. This version of the scorecard places greater emphasis on asset-quality measures than the original scorecard. This version also includes more gradations of risk within each of the three subcategories. The modified scorecard does maintain the net income, noncore funding, and equity elements of the original scorecard, but the weight placed on these measures has been reduced, as would be more appropriate for CAMELS 1- and 2-rated banks. Also, the equity measure has been changed to include loss reserves.[9]

[9] This equity measure, which includes loss reserves, performed better in statistical tests than equity alone. In addition, we believe that including loss reserves could create a disincentive to charge off loans purely to avoid higher nonperforming-asset scores.

Table 6

Scorecard
(Weightings Based on Well-Capitalized and Highly-Rated Institutions Only)

		Range of Scores	Maximum Score
Scoring Factor			
Nonaccrual Loans / Total Assets	< 0.5%	30	30
	= 0.5–1.0%	26	
	= 1.0–1.5%	23	
	= 1.5–2.0%	21	
	= 2.0–2.5%	20	
	> 2.5%	0	
Loans Past Due 90+ Days / Total Assets	< 0.5%	25	25
	= 0.5–1.0%	22	
	= 1.0–1.5%	20	
	= 1.5–2.0%	18	
	= 2.0–2.5%	13	
	> 2.5%	0	
ORE / Total Assets	< 0.5%	20	20
	= 0.5–1.0%	16	
	= 1.0–1.5%	14	
	= 1.5–2.0%	12	
	= 2.0–2.5%	11	
	> 2.5%	0	
Loans Past Due 30–89 Days / Total Assets	< 0.5%	14	14
	= 0.5–1%	12	
	= 1–1.5%	10	
	= 1.5–2%	9	
	= 2–2.5%	8	
	> 2.5%	0	
Pretax Net Operating Income / Average Assets	> 0.5%	7	7
	= 0–0.5%	4	
	< 0%	0	
Noncore Funding / Total Assets	<= 40%	3	3
	> 40%	0	
Equity & Reserves / Total Assets	> 7.0%	1	1
	<= 7.0%	0	
Total			**100**
Application of Scoring Framework			
If institution is 1A and total score is >= 97, classify as			1A1
If institution is 1A and total score is < 97 and >= 87, classify as			1A2
If institution is 1A and total score is < 87, classify as			1A3
Adjustment Factor if Noncore Funding / Total Assets > 40%			
Market Adjustment for Standard and Poor's AA– or Better		3	
Market Adjustment for Standard and Poor's A– to A+		1	
Market Adjustment for Standard and Poor's BBB+ or Worse		0	

Table 7 shows the distribution that would have resulted under this structure. It is clear that the distribution can shift significantly over different periods. Although 43 percent of the banks would have been classified in the best category at the end of 2002, only 24 percent would have been in this category in 1985. Thus, a certain amount of migration into and out of categories can be expected as the banking industry passes into and out of periods of stress.

Table 7

Distribution of Well-Capitalized and Highly-Rated Institutions Based on the Scorecard (1985–2002)

Year	Subcategories 1A1	1A2	1A3
2002	43.5%	38.6%	17.9%
2000	44.5	38.0	17.6
1995	45.5	42.2	12.3
1990	30.0	43.3	26.7
1985	23.8	41.4	34.9

Note: For each selected year, 1A institutions are scored on the basis of their reported financial ratios at year-end and are then placed into one of the three subcategories demarcated in the section of table 6 called "Application of Scoring Framework".

Relative Merits of Proposed Pricing Options

To recap, the pricing options presented here for banks (other than large banking organizations) include expanded use of supervisory ratings, use of statistical models (both in a continuous and discrete format), use of a combination of statistical models and supervisory ratings, and a scorecard that uses expert judgment in conjunction with a statistical model. How do these options fare relative to the desirable attributes outlined earlier?

Accuracy

We can compare accuracy, or the ability to differentiate risk, through the use of power curves for each of the deposit insurance pricing options discussed. Figure 3 is a power curve that represents how well each of the options performs in identifying failures. The horizontal axis of the figure shows the percentage of total institutions scored by each method. The institutions are sorted left to right, from those having the worst score (most likely to fail) to those having the best score. The vertical axis shows the cumulative percentage of total failures identified. The point identified on the figure shows that the first 10 percent of the institutions ranked according to SCOR values contained 62 percent of the total failures. The closer the curve is to the upper-left corner of the graph, the more accurate the particular method is at identifying failures. The diagonal line essentially represents a system with no predictive power, where the number of failures identified is proportional to the percentile of observations.

A failure identification score can be developed by measuring the area between an option's respective curve and the diagonal line. Based upon this score, the CAMELS-downgrade model has the most predictive power (37.05), followed by the failure-prediction model (35.66), the scorecard (28.49), and finally the supervisory-based structure (16.74).[10]

Simplicity

The supervisory ratings approach has an advantage over the other options when considering simplicity because of the level of familiarity with and acceptance of the CAMELS rating system. The CAMELS rating system is well understood and accepted by the banking industry, and broad agreement exists as to what each of the five ratings means in terms of a bank's condition. In contrast, the statistical models are more complicated than other methods. They also are less

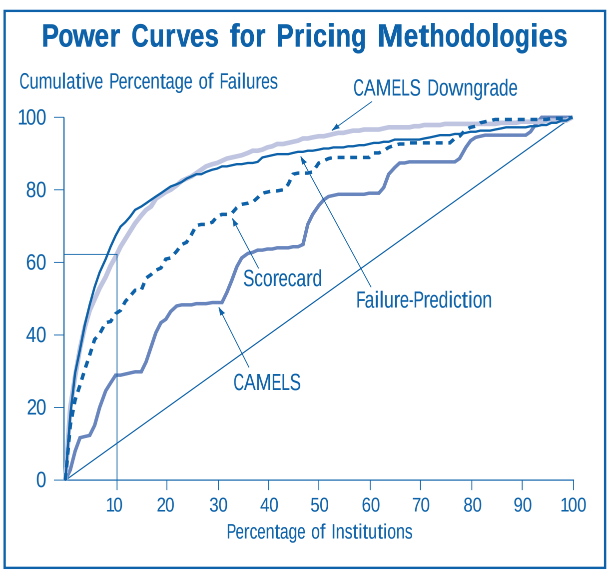
[10] This is not to say that supervisory ratings are inaccurate. Rather, CAMELS ratings provide a relatively broad measure of risk. They are good at separating healthier institutions from those showing more pronounced financial weaknesses, but are not specifically designed to differentiate among better-rated institutions. The statistical models, on the other hand, were developed to fit the historical failure and CAMELS downgrade data. In a sense, they are designed to excel at tests of historical accuracy (ex post). It is not clear that statistical models would fare as well relative to CAMELS ratings going forward (ex ante), where the task is to identify emerging risk factors that may or may not be reflected in historical experience.

transparent to insured institutions because the mechanics of the models are not observable. The scorecard represents an attempt to simplify the purely statistical approaches by combining an expert system with the failure-prediction model.

Flexibility

The supervisory ratings approach also holds an advantage over the other options in terms of flexibility. When examiners assign a CAMELS rating, they have access to and can analyze a wide range of data, including information about management, underwriting, and various intangible factors. Statistical models and scorecard can never completely reflect the current financial condition of a bank because they rely on Call Report information and because they cannot be tailored to reflect the unique aspects of individual banks. With a supervisory ratings approach, changes in virtually any factor predictive of bank failure, as well as improvements in supervisors' ability to measure risk exposure, would automatically be incorporated into the deposit insurance pricing system. To achieve flexibility, the statistical models and scorecard would need to be updated with some frequency to ensure that they continue to reflect the factors most closely associated with risk, thus making them more difficult to implement.

Figure 3



Appropriate Incentives

A supervisory approach also would best avoid perverse-incentive problems because examiners would verify on-site that operating results were achieved through safe and sound management practices. Supervisory and insurance ratings would therefore be closely aligned. Purely statistical approaches could create unintended incentive problems because they rely completely on Call Report data. The scorecard was designed to reduce the possibility of perverse incentives inherent in the statistical approach. The expert system incorporated into this approach would allow choices to be made regarding the factors used and threshold values to avoid these problems.

Fairness

Even though a goal of each of the pricing options is to treat banks equitably, almost always there will be cases in which the classification of a particular bank may be seen as unfair. The combined supervisory ratings and statistical models approach was developed, in part, to address a fairness issue. A purely statistical approach probably would result in some CAMELS 1-rated institutions paying higher premiums than some CAMELS 2-rated institutions. A combined approach could prevent this outcome.

Another way to evaluate fairness is in terms of an option's objectivity. The statistical models are the most objective since they rely completely on a bank's reported data. The scorecard would be the next most objective, followed by the supervisory ratings approach, which would rely a great deal upon examiners' judgment.

Pricing Options for New Banks

A separate pricing system is being considered for new institutions because of their special characteristics. Risks in new institutions often result from the fact that these institutions operate with unproven business plans in markets served by established competitors. The risks inherent in new institutions are not easily identified by the methods that can be useful for detecting risk in seasoned institutions. For instance, new institutions typically have high capital-to-asset ratios and low levels of problem assets compared with their seasoned-institution counterparts, yet new institutions have generally displayed a higher failure rate than seasoned institutions. As a result, pricing structures that rely on financial ratios would be less effective in identifying risk in new institutions.

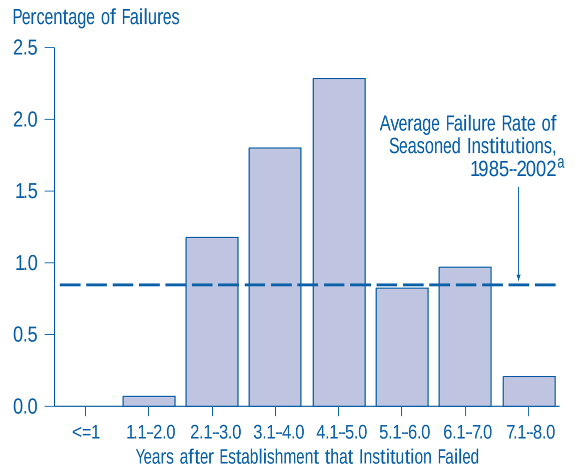
No consensus exists as to when a new bank takes on the characteristics of a seasoned institution. For purposes of deposit insurance pricing, we define new banks as those existing for five years or less. Figure 4 shows that at the end of this five-year period, failure rates of new institutions approach failure rates of seasoned institutions. Moreover, five years should allow new institutions enough time to confirm the viability of their business plans. Conversely, using a period longer than five years could discourage bank formation because of the relatively higher premiums to be paid by new institutions.

Two options for setting the assessment rate for new institutions are currently being considered. They are based on the premise that, although new banks should pay a risk premium that reflects their historical failure experience, the premium should not be so high as to discourage new firms from entering the industry. The two options are (1) automatically charge new institutions the highest rate paid by well-capitalized, highly-rated banks, or (2) charge new institutions a separate rate. The difference between the two lies in the maximum rate that could be charged to these new institutions.

In the second option—charging new institutions a separate rate—the assessment rate for these new institutions could be based on the historical risk profile of new institutions as a group, and the rate could be capped so that these new institutions would pay a rate lower than the rate paid by institutions that are less than well capitalized and not highly rated. As in option 1, new institutions that fall outside the best insurance assessment risk class would pay the same rate as other institutions in their particular class.

Figure 4

Years to Failure for Institutions Established between 1985 and 1994
(Failure Data through 2002)



[a] Seasoned institutions are institutions that have been in operation for at least five years.

Pricing Options for Large Banks

Large banks also have special characteristics, which may not be captured by the more traditional approaches to risk assessment. This unique status is explicitly acknowledged by FDICIA, which allows for separate pricing based on institution size. Although a pricing system that relies primarily on financial ratios derived from Call Report data may be suitable for small and medium-sized banking institutions, it may not be the best approach to identifying and equitably charging for risk in larger, more complex institutions.

In the long term, Basel II holds some promise for pricing large bank risk because it incorporates default probabilities derived from the institutions' own internal credit-risk models.[11] Such a system will not be possible before the Basel II capital guidelines are implemented (the scheduled date is 2007). In the meantime, information derived from the financial markets—alone or in combination with supervisory information—may provide a more accurate way to evaluate and price risk in large and complex organizations than an accounting-based system.

Developing a pricing system specifically for large and complex banking organizations will first require establishing criteria to select the institutions that would be subject to such an alternative system. The simplest and most commonly used criterion for delineating the group of large and complex banking organizations is asset size. Another criterion could be market capitalization, or a measure of complexity such as market participation or foreign operations.

Aside from the criterion used to define large banks, the ability to implement a pricing system that relies upon financial market data depends upon the availability of market data for these larger institutions. Equity data are generally available for most large banking companies. However, other market data, such as subordinated debt price quotes, are not available for several large banks.

Pricing Framework Based on Supervisory Ratings

A simple method for categorizing large, well-capitalized and highly-rated institutions according to risk is the method already proposed for small and medium-sized banks: creating two or more subgroups based on CAMELS ratings. Using supervisory ratings to set assessment rates for large institutions is appealing for several reasons. Large banking organizations are subject to frequent and thorough on-site review. Continuous supervision programs, which provide real-time and continuous evaluations of risk, have been established by the Office of the Comptroller of the Currency, the Federal Reserve Board, and the FDIC. Also, ratings assigned by regulators to large banks reflect information from a variety of sources, including the financial markets.

Pricing Framework Based on Market Measures of Risk

As the scale and complexity of the banking industry has increased, interest in using market information as a regulatory tool has grown. Regulators already use market signals extensively to monitor bank risk, and a variety of market indicators hold promise for pricing deposit insurance for large and complex institutions. These include price data such as stock price volatility and subordinated debt yield spreads; credit ratings assigned by companies such as Moody's, Standard and Poor's (S&P), and Fitch; and estimated default frequencies calculated using option pricing-type models such as the one developed by KMV Corporation.[12] A combination of these measures—and others that might prove suitable—could provide a more robust and balanced pricing tool for large banks than one based entirely on either

[11] As the requirements of Basel II are formalized and as institutions opt to adhere to them, we would expect the FDIC to incorporate information about the bank's internal credit rating systems, operational risk, and market risk into its pricing of deposit insurance.

[12] KMV's model calculates a company's probability of default from its stock price volatility, current capital structure, and value of its assets.

supervisory ratings or financial ratios. A disadvantage of relying entirely on market measures is that the insurer would forgo the benefits of information gleaned by examiners with access to confidential information.

A key issue is whether the data used for insurance pricing should originate at the bank or the parent holding-company level. In general, data related to the depository institution are of greater value to the insurer, since they reflect the consensus opinions of investors about the condition and performance of the entity having the most direct access to the federal safety net. Data related to the depository institution are all the more important in light of the increasing diversification of financial holding companies into business lines unrelated to banking; market information about a parent company may not accurately reflect the performance of an insured subsidiary.

Unfortunately, market data are often unavailable at the insured-institution level. The equity and debt instruments that would provide information useful for deposit insurance pricing are more typically issued by the parent holding companies of banks. This practice may compel the use of holding-company data for pricing.

Measuring the Predictive Ability of Market Factors

Because so few large institutions have failed, insufficient data are available to establish statistical relationships between the probability of bank failure and market measures in the same way failure was correlated with Call Report data to develop the failure-prediction model. To establish the usefulness of market measures as predictive factors, therefore, we tested three market measures against supervisory downgrades from CAMELS 1 or 2 to CAMELS 3 or lower over the period from 1987 through 1999 for the largest 25 banks as of year-end 1999.

Figure 5 shows the degree to which stock price volatility has predicted downgrades. We calculated a coefficient of variation for stock price (as a measure of volatility) and grouped the institutions by high, medium, or low volatility. The bars in the figure show the percentage of banks in each category that were downgraded to a composite 3 rating or worse within two years of our calculation. The results show a relationship between stock volatility and supervisory downgrades, indicating that stock price volatility may be an effective way to differentiate institutions for pricing purposes.

Figure 6 shows how well S&P credit ratings perform in predicting CAMELS composite downgrades. These aggregate results show a certain degree of differentiation between higher and lower investment-grade ratings, and a significant differentiation between investment- and noninvestment-grade ratings.

Figure 5

Aggregate CAMELS Downgrades in Relation to Stock Price Volatility for Top 25 Institutions
(Coefficient of Variation, 1987–1999)

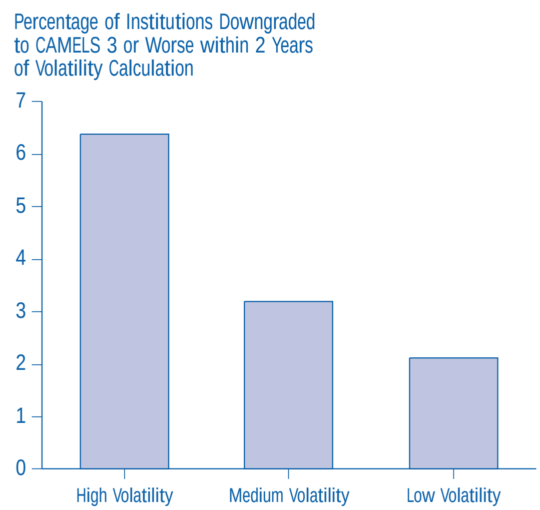


Figure 7 shows how KMV-estimated default frequencies perform in predicting downgrades during subsequent two-year periods. Again, a correlation is evident between the market measure and higher probabilities of downgrades.

Methodology for Assigning Scores

A large-bank pricing system using market measures could be constructed in a number of ways. One would be to use an individual measure—for example, credit ratings—on its own. Currently, S&P credit ratings for the parent companies of the 50 largest insured institutions range from AA– to BBB–. Given this fairly wide distribution, assessment rates could be assigned either to each credit rating category individually or to larger groups made up of more than one rating category. Alternatively, a combination of market measures could be weighted and summed to produce a single score per institution.

Pricing Framework Based on Combination of Supervisory and Market Measures

Another way to create subcategories in the best insurance group would be to use supervisory ratings in combination with a select set of market measures of bank risk. Such a system could take the form of either an integrated system in which supervisory ratings and market measures were combined and equally represented or a system in which market measures would serve as trip wires to adjust insurance classifications based mainly on supervisory ratings. An integrated system would require a method of weighting the various factors—composite ratings and market measures—to produce a single score.

A pricing system with trip wires might incorporate any of the market measures mentioned above (or others, such as price-to-book ratios) to adjust institution scores after the institutions had initially been categorized by supervisory ratings. For example, banks might be placed into separate CAMELS 1 and CAMELS 2 categories, and CAMELS 2-rated institutions that had relatively poor credit ratings might then be relegated to a third category. In table 8, the composite 2-rated group is subdivided into two categories: those with S&P credit ratings of A– or better (the 1A2 group) and those with ratings worse than A– (the 1A3 group). Similarly, high subordinated debt yield spreads, high stock price volatility, or low price-to-book ratios might serve as the secondary means of differentiation.

Figure 6

Aggregate CAMELS Downgrades in Relation to Credit Rating for Top 25 Institutions
(Coefficient of Variation, 1987–1999)

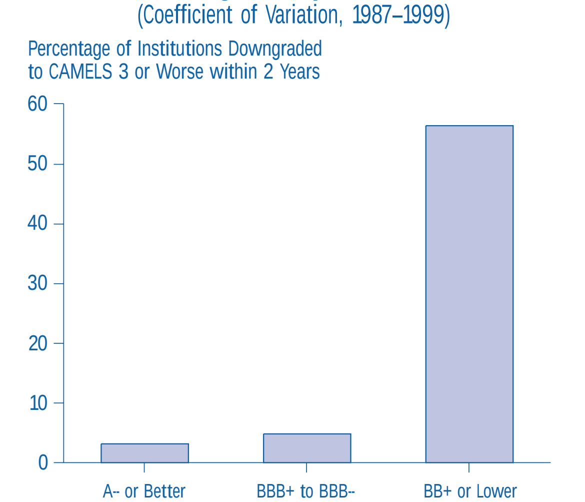
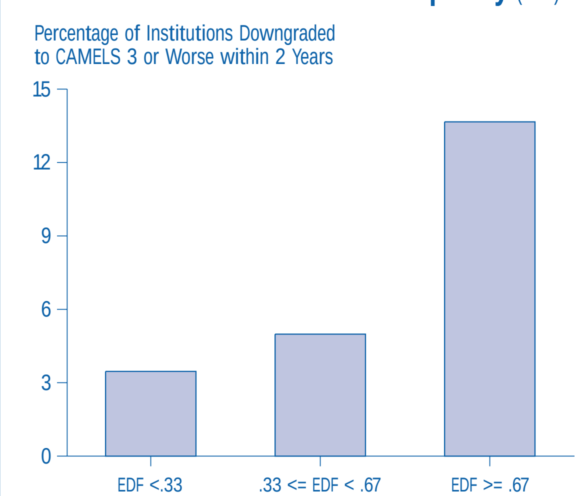


Figure 7

Aggregate CAMELS Downgrades in Relation to KMV for Estimated Default Frequency (EDF)



Pricing with a Scorecard

The disadvantages of using a financial ratio-based pricing system for large banks are discussed above. However, it may be possible to modify the scorecard approach in ways that would eliminate unintended adverse effects on large banks. Larger institutions tend to be penalized by the noncore funding component of the scorecard because they often operate with higher levels of wholesale funding and lower levels of capital than the smaller institutions that compose the bulk of observations used to calibrate the scorecard risk weights. To compensate, selected market measures could be incorporated into the scorecard either to replace certain of its elements as measures of risk or to offset elements that unduly penalize large banks. One approach we explored is the use of credit ratings.

For example, the lowest section of table 6 (see page 9) shows the part of a revised scorecard that includes an adjustment for institutions' S&P credit ratings. The rating adjustment relates to a bank's noncore funding score: banks that have ratings of A– or better receive an upward scoring adjustment to reflect their enhanced ability to obtain capital in the debt markets. Table 9 shows the effect of the rating adjustment on the distribution of large-bank rankings based on the modified scorecard. The rating adjustment results in an increase in the percentage of large banks placed in the risk category 1A1, though this percentage remains below the percentage of small banks in the 1A1 category. The percentage gap between large and small 1A1 institutions may reflect the relatively stronger asset-quality measurements for smaller institutions.

Table 8

Options for Pricing Large Institutions Using CAMELS Ratings with Credit Ratings (Year-End 2002)

	Subcategories		
	1A1	**1A2**	**1A3**
Composite 1 Rated	15 32.6%		
Composite 2 Rated *and* Credit Rating A– or better		24 52.2%	
Composite 2 Rated *and* Credit Rating Worse than A–			7 15.2%

Note: The figures in the cells refer to the number and percentage of the 46 largest institutions for which S&P ratings are available.

Table 9

Distribution of Well-Capitalized and Highly-Rated Institutions by Size Based on the Scorecard (Year-End 2002)

	Subcategories		
	1A1	**1A2**	**1A3**
Small Banks	43.5%	38.5%	18.0%
Large Banks[a] with Noncore Funding Adjustment	36.7	57.1	6.1
Large Banks[a] without Noncore Funding Adjustment	32.7	59.2	8.2

Note: The scorecard-derived scores (see table 6) produce the distribution shown here. In the scorecard, the adjustment for noncore funding rewards institutions rated by S&P as AA– or better with a 3-point upward adjustment, and institutions rated A– to A+ with a 1-point upward adjustment.

[a] Large banks are the top 50 banks by asset size.

Conclusion

All the options discussed in this article involve trade-offs among the desirable attributes of a deposit insurance pricing system. As applied to historical data, the statistical approaches tend to provide greater risk differentiation than the supervisory ratings approach but also tend to be more complex, more difficult to implement, and more likely to create unintended perverse incentives. The scorecard has an advantage over pure statistical models in terms of simplicity, flexibility, and incentives, but it is less accurate.

The combined statistical and supervisory approach was presented as an option that can ensure that CAMELS 1-rated institutions never pay more than CAMELS 2-rated institutions. Also, the combined approach supplements the informational content of the CAMELS ratings with the more recent information reported in Call Reports. However, the combined approach does not eliminate all the disadvantages of either of the two pure approaches. For example, if a combined methodology breaks the well-capitalized and highly-rated group of institutions into three or four subcategories, there is still the potential for cliff effects—small changes in a measured variable that produce large changes in the deposit insurance premium. In addition, a combined system is more complex than a system based on CAMELS ratings alone.

Nonetheless, combining the statistical and supervisory approaches can mitigate several of the concerns relating to either approach in isolation. The combined approaches and perhaps the scorecard approach provide the opportunity to make practical trade-offs and achieve the right balance among desirable attributes and policy objectives.

Separate deposit insurance pricing options were presented for new banks to address their special characteristics. Additional options also were presented for large banks that incorporate market data, which may better identify risk in larger, more complex institutions. Ultimately, the selection of one or another approach will reflect a particular weighting of the desirable attributes and a judgment regarding the approach that achieves the best balance.

Evaluating the Vulnerability of Banks and Thrifts to a Real Estate Crisis

*Charles Collier, Sean Forbush, and Daniel A. Nuxoll**

As part of its extensive off-site monitoring efforts, the Federal Deposit Insurance Corporation (FDIC) has evaluated banks' and thrifts' vulnerability to the stress of a real estate crisis similar to the crisis that occurred in New England in the early 1990s.[1] Asking what would happen to banks and thrifts today if the real estate market were to experience a downturn similar to the one in New England a decade ago, we developed the history of the collapse of the New England real estate market into a stress test—the Real Estate Stress Test (REST)—that produces ratings comparable to the CAMELS ratings.[2] The REST ratings indicate the severity of the exposure to real estate and therefore identify institutions that appear vulnerable to real estate problems. The ratings direct the attention of examiners to particular institutions and indicate that the FDIC should be especially concerned about the management of real estate lending at these institutions. Poor practices there could expose the FDIC to substantial losses.

In addition, REST is able to identify particular areas of the country where a high fraction of the banks and thrifts are vulnerable—areas where the real estate markets might be of concern to bank examiners. Although these markets may be healthy at the moment, the extent of bank lending in them means that the FDIC must pay particular attention to conditions there.

The results of our research with REST indicate that the institutions most vulnerable to real estate crises today are headquartered in the West and a

*All the authors are on the staff of the Federal Deposit Insurance Corporation (FDIC). Charles Collier and Sean Forbush are with the Division of Supervision and Consumer Protection (DSC), Collier as chief of the Information Management Section and Forbush as a senior financial analyst. Daniel Nuxoll is with the Division of Insurance and Research (DIR) as a senior economist.
This article reports the results of a close collaboration among numerous people in both the DSC and the DIR. In addition, the staff of the FDIC's San Francisco Regional Office encouraged the project and provided the authors with helpful comments.
The opinions expressed here are those of the authors and do not necessarily reflect the views of the FDIC.

[1] See Collier et al. (2003) for a more general discussion of the objectives and methods of the FDIC's off-site models.

[2] CAMELS ratings are based on examiners' assessments of **C**apital, **A**sset quality, **M**anagement, **E**arnings, **L**iquidity, and market **S**ensitivity. The ratings range from 1 to 5, with 1 being the best. Banks and thrifts with a rating of 1 or 2 are considered sound, whereas supervisors have definite concerns about institutions with a rating of 3. Institutions with a rating of 4 or 5 are considered problem banks. The Sensitivity rating was added only in 1997, so strictly speaking, ratings before that year are CAMEL ratings. This article uses "CAMELS" throughout, despite the anachronism.

handful of southern cities.[3] The real estate markets in these locations are currently healthy, but because banks—and by extension the FDIC—have substantial exposure to these markets, bank supervisors need to be especially alert to any indication of problems there.

We also find that the most critical risk factor is construction lending, a finding that confirms the conventional wisdom that construction lending is particularly risky. Many accounts of the savings and loan crisis of the late 1980s and early 1990s discuss commercial and residential construction projects that went awry.[4]

Because the stress test was developed on data from New England, it may well reflect the distinctive characteristics of events in that region. However, when REST was backtested on data from Southern California in the late 1980s and early 1990s, it was successful in identifying institutions that later had problems. More importantly, REST was also successful in identifying troubled banks in parts of the country where real estate downturns were moderate. These successes suggest that even if a repetition of the severe problems of New England or Southern California is very unlikely, REST can still help identify banks that might suffer difficulties during less severe real estate downturns.

The REST model should not be interpreted as a condemnation of construction lending. The model does, however, emphasize that risk control is especially important for these loans. The success of a construction loan depends on the future, not the present, of the real estate market, so construction lending is intrinsically more risky than forms of lending that are secured by liens on real property.

The obvious question is why one should focus on New England. There are three reasons. First, problems among the banks in New England can be traced directly to the real estate market.[5] Second, the number of banks in the region was large enough that statistical models can be estimated relatively easily. Third, the New England experience is hardly unique. As the FDIC (1997) documents, commercial real estate was a factor in several distinct sets of banking problems during the 1980s and early 1990s.[6] In addition, commercial real estate has been a factor in bank crises in a number of other countries.[7] Thus, events in New England constitute a relatively clear case of a problem that is endemic to banking.

Importantly, REST uses Call Report data, so it cannot evaluate pricing, terms, or underwriting—factors critical to controlling the risk of real estate lending. Moreover, REST does not estimate the condition of the real estate market in any region, state, or metropolitan statistical area (MSA); it identifies markets where banks are exposed to potential real estate problems, not markets where such problems actually exist. What the REST model can do is identify the banks that are at most risk in the event real estate problems should occur. In so doing, it sharpens the focus of questions about risk control and real estate markets and therefore makes an important contribution to the FDIC's off-site monitoring.

This article explains how the model was built with the use of New England data and was tested with the use of data from other historical real estate crises. The REST results for December 2002 are presented and analyzed, and recent trends—both nationally and for selected states—are discussed.

[3] Clearly, our project is most directly related to the FDIC's function as an insurer, not a supervisor. Consequently, this article discusses all banks and thrifts, whether or not they are supervised by the FDIC.
It must also be observed that banks are identified by their headquarters. Consequently, for purposes of this stress test, the Bank of America is located in Charlotte, N.C., although the vast majority of its business is outside the Charlotte metropolitan statistical area and outside the state of North Carolina. However, the number of megabanks is relatively small, and few of the banks in our project have many operations that are outside a small area.

[4] A number of popular accounts–for example, see Mayer (1990), chapter 5–report that Edwin Gray, the chairman of the Federal Home Loan Bank Board from 1983 to 1987, became aware of the depth of the S&L crisis while watching a videotape of abandoned projects in the Dallas area.

[5] See FDIC (1997), chapter 10, for a discussion of this issue. In contrast, the Texas banking crisis during the late 1980s and early 1990s was caused only partly by commercial real estate.

[6] Ibid., especially chapter 3.

[7] See Herring and Wachter (1999).

Method of Examining New England

The central question for the team that built the REST model was whether any model could detect those healthy banks that would be in most danger during periods when real estate became a problem. To answer this question, we examined the New England real estate crisis of the early 1990s. In 1987, the economy and the banking industry in New England could have been described as vibrant, but by 1990 the problems were obvious.[8] The first stage in developing the REST model involved comparing the banks in New England in 1987 with the banks there in 1990. All the banks were healthy in 1987, but by 1990 a substantial fraction of them were troubled. Our analysis used statistical procedures and data from 1987 to find the traits common to the institutions that later had severe difficulties. This approach seeks to answer the question whether as early as 1987 one could have identified the riskiest banks in New England.

Because the purpose of our project was to evaluate banks' ability to withstand a crisis such as the one in New England in 1991–1993, banks that had a special function or were somehow atypical were eliminated from the analysis. Banks considered atypical were those that had equity-to-asset ratios greater than 30 percent or loan-to-asset ratios less than 25 percent. A total of 13 special-purpose or atypical (or new) banks were eliminated from the December 1987 sample.[9]

In addition, consolidation just before the crisis had to be taken into account. In December 1987, 289 New England banks filed Call Reports, but in December 1990 the number had shrunk to 255. Much of the consolidation appears to have been achieved by mergers of different banks owned by the same holding company. Regardless of the reason for the consolidation, the performance of the bank resulting from a merger was undoubtedly affected by the characteristics of the banks absorbed in the merger. Consequently, this project used data adjusted for mergers.[10]

Finally, because growth rates between 1985 and 1987 were included in the model, only banks that had been in existence for five years (1985–1990) were part of the sample.

The sample contained a total of 203 banks.[11]

In the first stage—comparing the conditions and balance sheets of banks at the end of 1987 with the same banks' conditions and balance sheets at the end of 1990—the model considers 12 variables as measures of health at the end of 1990 (previous work has shown that these variables are closely related to CAMELS ratings, and the FDIC has developed a Statistical CAMELS Off-site Rating [SCOR] model using them).[12] The 1987 data include the same 12 variables as well as 12

[8] We could have used data from years other than 1987 and 1990 to develop the REST model, but for a terminal date, 1990 is the obvious choice. The problems in New England were not that apparent until 1990, yet in 1991 a significant number of banks failed. We are especially interested in banks that are so troubled they eventually fail; thus, a later terminal date would ignore some important information. The start date of 1987 corresponds closely to the peak in the New England economy, but 1986 or 1988 could equally well have been used. Experiments indicate that the REST results would have been similar for any of those three years.

[9] Also excluded was a Connecticut bank that at the end of 1988 apparently sold its regular banking operations and continued as a special-purpose institution.

[10] To adjust the data, we combined the data for separate institutions that later merged. For example, if two banks merged in January 1988, the 1987 data for the resulting bank would be the combined balance sheets and income statements for the two banks as of December 1987.

[11] Our discussion of New England does not refer to thrifts because the savings banks were excluded from the sample. During this period, savings banks filed a slightly different Call Report from the one filed by commercial banks, so some data provided by commercial banks are missing for savings banks.. More importantly, during this period many mutual savings banks converted to stockholder-owned savings banks, and after conversion, these institutions behaved quite differently. See FDIC (1997). The development of the stress test assumes that the institutions in the sample had a generally stable strategy, and clearly many of the savings banks in New England did not. Our discussion of Southern California does not include thrifts because before 1991, data on thrifts in that region are limited.

[12] See Collier et al. (2003). A model could be developed that would forecast CAMELS ratings directly. However, the deterioration among banks in New England was extremely sudden, and CAMELS ratings change only after an examination (or, occasionally, after an off-site review). CAMELS ratings at the end of 1990 probably do not reflect the extent of the problems in New England because examiners were overwhelmed and had not changed the ratings at some troubled institutions. We developed a model to forecast CAMELS ratings directly, and although it identified the same types of institutions as the REST model, in backtests it was found to be slightly less accurate than the REST model.

variables that measure (as a fraction of assets) the types of loans made by the bank, a variable that measures the bank's growth rate between 1985 and 1987, and a variable that measures the bank's size in 1987.

The basic results appear in tables 1 and 2. The reason the two tables differ is that only 3 of the 12 SCOR variables (equity, provisions for loan losses, and net income) can be less than zero, while another of the variables (loans and long-term securities) can be zero in principle but in fact was substantially greater than zero for the whole sample. These 4 variables were handled by the usual regression technique—ordinary least squares (OLS). Table 1 reports the results for these 4 variables.

Table 1

The REST Model: Coefficients Estimated with OLS

	Equity	Provisions	Income before Taxes	Loans
Intercept	6.111	–3.830	4.933	22.523
Log assets	–0.315	0.283	–0.245	
Growth	–0.020	0.008	–0.009	0.046
Lagged Variables				
Equity	0.535			
Loan-loss reserves	1.346			
Loans past due 30–89 days		0.253	–0.354	–0.960
Loans past due 90+ days	–0.804			
Nonaccrual loans		–0.755	1.029	
Other real estate				
Charge-offs		1.084		
Provisions for loan loss				
Income before taxes	0.815		0.378	
Noncore liabilities			–0.032	
Liquid assets	0.032			
Loans and long-term securities				0.689
Loan Types				
Agriculture loans				
C&I loans	–0.040	0.054	–0.075	
Credit card loans	0.059		0.072	
Other consumer loans				
Loans to depositories				
Municipal loans				
Agricultural real estate loans				0.858
Construction loans	–0.229	0.217	–0.366	
Multifamily-housing loans				
Nonresidential real estate loans				
1–4 family mortgages				
Leases		–0.214	0.376	
R^2	0.5212	0.3702	0.4474	0.5834
F-Statistic	0.676	0.868	0.557	0.611
Degrees of freedom	16,176	18,176	16,176	22,176

Note: The data are for 203 New England banks. The independent variables are from December 1987 and the dependent variables are from December 1990. Charge-offs, provisions, income, and growth are all based on merger-adjusted data.

The other 8 SCOR variables cannot, in principle, be less than zero. These variables—loan-loss reserves, loans past due 30–89 days, past due 90+ days, nonaccruals, other real estate, charge-offs, volatile liabilities, and liquid assets—were fit with a Tobit model.[13] Table 2 reports the results for these variables. For a number of the 8 variables, the results do not differ appreciably from OLS because, as reported in table 2, there were very few zero values.[14]

As mentioned above, the independent variables for the REST model include all the 1987 values for SCOR variables, 12 categories of loans as a fraction of assets, asset growth, and bank size.

The SCOR variables represent the condition of the bank in 1987. In fact, the condition of the bank results partly from the characteristics of the bank, so these 12 variables are proxies for the characteristics of the bank. For example, one cannot directly observe the quality of a bank's underwriting, but presumably tighter underwriting results in fewer past-due loans—so the data on loans past due 30–89 days can be seen as a proxy for underwriting standards.

The loan-type variables are important because the New England crisis was a real estate crisis. Our project included data on

[13] Other real estate consists mostly of real estate that banks own because of foreclosures. Charge-offs are gross, not net, so they cannot be less than zero.

[14] In fact, all banks had some loans past due 30–89 days, but the OLS estimates differ from Tobit because of a handful of values that are close to zero. Tobit considers the possibility that these values are greater than zero by chance.

several types of real estate loans (1–4 family residential, multifamily housing, agricultural, construction and development, and other nonresidential) as well as other loans (unsecured commercial, to municipalities, to depository institutions, credit card, other consumer, agricultural production). Presumably banks that held large amounts of real estate loans would be the ones most severely affected by the crisis.

Asset growth between 1985 and 1987 was included because rapidly growing banks are considered

Table 2

The REST Model: Coefficients Estimated with Tobit

	Reserves	Past Dues 30–89	Past Dues 90+	Nonaccrual	Other Real Estate	Charge-offs	Noncore Liabilities	Liquid Assets
Intercept	–2.385	5.756	2.595	–5.627	–0.706	–3.026	–8.015	24.764
Log assets	0.224	–0.278	–0.113	0.444		0.181	0.803	0.963
Growth	0.003				0.006	0.007	–0.042	–0.052
Lagged Variables								
Equity				0.130		0.045	0.597	
Loan-loss reserves		0.554	–0.582					
Loans past due 30–89 days		0.646		0.373	0.440	0.256	–0.930	
Loans past due 90+ days			0.455					
Nonaccrual loans		0.790	0.416				–2.621	
Other real estate			–0.869	–1.407				
Charge-offs		–2.002				0.712	–11.673	
Provisions for loan loss							11.304	
Income before taxes								
Noncore liabilities		0.028			0.026		0.696	
Liquid assets		–0.036						0.428
Loans and long-term securities								–0.247
Loan Types								
Agricultural loans								
C&I loans	0.031			0.044		0.019		
Credit card loans							0.105	
Other consumer loans	0.025	0.044						
Loans to depositories	0.189						–0.819	
Municipal loans		–0.150	–0.121		–0.240			
Agricultural real estate loans						–0.076	0.820	
Construction loans	0.113	0.077	0.071	0.250	0.155	0.088		–0.197
Multifamily-housing loans		0.144			0.132			
Nonresidential. real estate loans					0.069	0.025		–0.125
1–4 family mortgages			–0.012					
Leases								
Pseudo-R^2	0.3689	0.4597	0.1306	0.3588	0.3219	0.4380	0.7679	0.5862
Chi-Squared Statistic	12.45	7.67	9.43	16.76	17.05	12.44	9.00	17.51
Degrees of Freedom	20	15	18	20	19	17	15	20
Zero values	0	0	21	19	14	1	3	0

Note: The data are for 203 New England banks. The independent variables are from December 1987 and the dependent variables are from December 1990. Charge-offs, provisions, income, and growth are all based on merger-adjusted data.

especially risky. Total assets were included in the model because it is usually thought that larger banks can more easily diversify risk away.[15]

All estimations were done with a stepwise procedure.[16] This method starts with all 26 variables (12 SCOR variables, 12 loan-type variables, asset growth, and size) and eliminates those that are not statistically significant. The stepwise method was necessary because some variables have coefficients that are very large but statistically insignificant. Although inclusion of these variables improves the in-sample fit of the model, it does so only very slightly. If the coefficients are large, however, inclusion of these variables in out-of-sample forecasting would almost certainly have an effect on the forecasts despite the complete absence of statistical evidence that these variables matter at all. Their elimination made very little difference to the fit of the model.

New England Results

As noted above, all the results were estimated with a stepwise procedure, and this procedure did not result in a significantly worse fit than if all the variables had been used. In general, the two sets of estimates are completely consistent with each other. In fact, most of the coefficients estimated with a stepwise procedure are very similar to those estimated when all the variables are used.[17]

Although alternative methods produced similar estimates, one should be cautious about interpreting these results. For example, one cannot conclude that the ratio of loans and long-term securities to assets did not affect asset quality, although that variable has a zero coefficient in all the asset-quality equations (loans past due 30–89 days, loans past due 90+ days, nonaccrual loans, other real estate, charge-offs, and provisions for loan loss). The effect might be small or inconsistent. Statistical tests reveal correlation, not causation.[18] When the correlation is strong and consistent with theory, however, there is good reason to take statistical results seriously. With that in mind, one should note several features of the results.

First, this approach captures much of the variation between banks. Near the bottom of both table 1 and table 2 there is a line reporting that R^2 is between 0.30 and 0.60 for most of the results.[19] This means that 1987 data can account for about 30–60 percent of the differences between banks in 1990. The major exception to this result is that the variable "loans past due 90+ days" has an R^2 of only 0.1306.[20]

Second, most variables are mean-reverting. That is, the banks that were exceptional in 1987 tended to resemble the average (mean) bank more closely by 1990. The coefficients on the lagged variables show this effect. For example, consider the effect of lagged equity on equity. From table 1, the estimated coefficient is 0.535. This means that an extra 1 percent equity would lead to an extra 0.535 percent equity in 1990. Importantly, the coefficient is between 0 and 1, indicating that banks with unusually high levels of equity in 1987 still had unusually high levels of equity in 1990, but

[15] The number actually used is the logarithm of total assets.

[16] The statistical software SAS supports a stepwise method for OLS but not for Tobit. The variables with the Tobit specification were also estimated with stepwise OLS and with a full Tobit model (one that includes all 26 variables). The variables that were insignificant in both the stepwise OLS and the full Tobit specification were dropped. The Tobit was reestimated, and the more insignificant variables were dropped. In the final estimation, all variables were significant at least at the 15 percent level.

[17] It should be noted that because these equations were estimated with a stepwise procedure, the coefficients and t-statistics cannot be interpreted in the textbook manner. However, the estimated coefficients and t-statistics are very similar when all the variables are included.

[18] The stepwise procedure complicates the usual warning about reasoning from correlation to causality. The coefficient on a correlated variable might well incorporate the effect of an omitted variable.

[19] The numbers reported for the Tobit are pseudo-R^2s. They are calculated in a manner analogous to the manner in which OLS R^2s are calculated, except that with the Tobit numbers the calculation allows for the fact that the variables can never be less than zero.

[20] The test statistics for the hypothesis that the omitted variables have a zero coefficient are also included. By way of comparison, the 5 percent significance level for a Chi-squared statistic with 15 degrees of freedom is 25.00, while the comparable F-statistic with 20 and 200 degrees of freedom is 1.62. However, because the model was fitted with a stepwise procedure, the statistics in the tables are not useful for classical hypothesis testing. They merely indicate that excluding the variables has very little effect on the fit of the model.

other things being equal, differences in equity levels shrank during those three years. An inspection of tables 1 and 2 shows that the only variables without a strong mean-reverting component are provisions, reserves, nonaccrual loans, and other real estate.

This observation suggests that most of the SCOR variables reflect something fundamental about the operations of a bank. Banks with higher than average levels of loans past due 30–89 days tend to have higher than average levels even three years later. Conceivably, high levels of past-due loans may reflect a less cautious underwriting philosophy.

This interpretation of the SCOR variables is supported by the data. For example, high levels of loans past due 30–89 days might be considered a sign that the bank is more willing to take risks. In fact, high levels of loans past due 30–89 days in 1987 are associated with lower net income and more nonaccrual loans, more other real estate, more charge-offs, and more provisions in 1990.

The third feature of our New England REST results is that most of the loan-type variables have the expected coefficients. High levels of commercial real estate loans in 1987 were associated with poor performance in 1990. It should be noted that construction and development loans in particular were problems for New England banks. Although other types of commercial real estate (nonresidential real estate and multifamily housing) were associated with problems, construction loans were the major problem: they were significant in almost every regression, and they generally had a larger effect than other types of commercial real estate loans. High levels of commercial and industrial (C&I) loans and other consumer loans also seem to have been a risk factor. Credit card loans were not a special problem, and loans to municipalities helped shield banks from the downturn.

Fourth, high asset growth between 1985 and 1987 also resulted in poor performance by 1990. The signs on log assets are consistent with the theory that larger institutions were more diversified and more aggressive in facing their problems in 1990. Large institutions had fewer past-due loans; on the other hand, they had more nonaccrual loans, reserves, charge-offs, and provisions. They also had lower net income, but that result seems to be driven completely by the higher provisions.

There are some other interesting features of the results. Banks with high net income in 1987 tended to have higher equity in 1990. Banks with high levels of reserves in 1987 performed better in 1990. This last finding is consistent with the interpretation that more-conservative banks tend to recognize losses more quickly and reserve against them. This interpretation, in turn, is consistent with the observation that charge-offs in 1987 are negatively correlated with loans past due 30–89 days in 1990. Banks that relied on noncore liabilities also tended to have more difficulties in 1990 (lower income, more past dues 30–89 days, and more other real estate). This result is consistent with the notion that banks that use noncore liabilities may be more aggressive and take more risks.

And there are some anomalies. High levels of other real estate in 1987 are correlated with low levels of loans past due 90+ days and nonaccrual loans in 1990. This might reflect differences in workout policies.

Out-of-Sample Testing

Although these results are intrinsically interesting as an analysis of past events in New England, the goal of our project was to develop a forecasting tool that could identify banks most likely to have difficulties during future real estate downturns. To test whether REST had forecasting power, we applied it to other real estate crises. Because these tests involved banks that were not in the sample used to build the model, they are called "out-of-sample" tests.

Southern California experienced a real estate crisis at about the same time as New England. To test the validity of the New England results, we forecasted 1991 SCOR ratios on the basis of 1988 data for Southern California banks. The banks included all California banks overseen by the FDIC's Los Angeles East, Los Angeles West, and Orange County field offices. Again, all institutions with loan-to-asset ratios less than 25 percent or equity-to-asset ratios greater than 30 percent were excluded. The sample contained 242 banks, 173 of which had a composite CAMELS rating of 1 or 2 as of year-end 1988.

The banks in California differed from those in New England in a number of ways. First, California banks in 1988 were generally in worse shape than New England banks in 1987. None of the New England banks in the sample had a CAMELS rating worse than 3 at year-end 1987, but 20 (8.3 percent) of the Southern California banks were rated 4 at year-end 1988. Second, the shock in the New England economy and therefore to the region's banks was both shorter and more severe. Of the 203 New England banks, 33—16.3 percent—failed, a percentage slightly higher than the percentage in California (33 of 240, or 13.8 percent). Moreover, in New England the bulk of the failures (29) were concentrated in a two-year period (1991 and 1992), whereas in Southern California the failures were spread out over three years (1992–1994). Third, structural differences between the two regions' banking industries were significant: California had permitted statewide branching for decades, whereas the banking industry in New England was more segmented.

The stress test did not do particularly well at forecasting individual ratios. For example, the model was not able to identify those banks that experienced large increases in nonaccrual loans. This is not too surprising because management's decisions about handling problems determine how the problems affect the bank's balance sheet and income statement, and if bank management delays dealing with real estate problems, the bank will tend to have higher other real estate owned or nonaccrual loans. If management deals with the problems aggressively, those same problems may affect the bank's provisions, charge-offs, income, and capital. Even a perfect model cannot forecast how management will deal with problems.

However, bank supervisors do not evaluate banks in terms of individual ratios but in terms of the overall condition of the bank. Consequently, the major issue is whether the stress test can forecast bank condition. The SCOR model can be used to translate the 12 SCOR ratios into a forecasted CAMELS rating.[21] These ratings are the REST ratings.

Table 3 compares 1988 REST ratings with CAMELS ratings and with failures between 1992 and 1995. All the banks used for compiling table 3 were 1 or 2 rated as of December 1988, and all banks survived until at least December 1991. If the bank failed between 1992 and 1995, the bank is identified as a failure. Otherwise, the bank's reported CAMELS rating is the worst rating it received between 1992 and 1995.[22]

Several considerations underlie this approach. First, the ultimate concern of supervision is troubled banks; hence, one should concentrate on the worst ratings. Second, banks that are rated 3 or worse have already been identified as potential problems, and the critical question is which banks currently regarded as sound are likely to develop problems.[23] Third, as noted above, events in Southern California evolved over a number of years. Problems at a bank that were obvious at the end of 1993 might not have been evident at the end of 1991. Using the worst rating during the crisis years 1992–1995 avoids the issue of timing. This method considers the banks that encountered difficulties, regardless of when the problems actually occurred.

[21] Our project focuses on the information that could have been known at the time. Consequently, the REST ratings are computed with the same coefficients that could have been used to produce the December 1988 SCOR ratings. There is one complication: the coefficients were estimated using revised Call Report data and a complete set of examination ratings. Neither would have been available if someone had estimated the SCOR model in 1989.

[22] Three banks are excluded because although they survived until December 1991, they merged before they were examined. The mergers were not assisted; that is, the banks did not fail.

[23] The results are not materially different if one includes banks that were rated 3, 4, or 5 as of 1988.

Forecasting models have two types of error: they fail to identify the banks that are downgraded (Type I error), and they identify banks that are not downgraded (Type II error).[24] This article analyzes the number of banks that the model correctly identified, so it refers to Type I accuracy and Type II accuracy. The emphasis is on problem banks (banks with a CAMELS rating of 4 or 5) and failures. Failures cost the FDIC money to resolve the bank, and problem banks are in danger of failing and take considerably more supervisory resources.

Panel A of table 3 shows the raw numbers, while panel B reports Type I accuracy and panel C reports Type II accuracy. Ideally all banks that failed would have REST ratings worse than 4.5 (100 percent Type I accuracy), and the "failed" line in panel B would have a 100 percent in the last column. In addition, ideally all the banks with REST ratings of 4.5 or worse would eventually have a CAMELS rating of 5 or would fail (100 percent Type II accuracy). In that case, the column for REST ratings greater than 4.5 in panel C would have numbers that sum to 100 percent in the lines for CAMELS ratings of 4 or 5 or for failures.

[24] For a more extended explanation of Type I and Type II errors, see Collier et al. (2003).

Table 3

Performance of Stress Test in Southern California

Worst Exam Rating	REST Rating 1.0–2.5	2.5–3.5	3.5–4.5	4.5–5.0	Total
A. Number of Banks					
1 or 2	12	12	18	4	46
3	12	17	19	5	53
4 or 5	4	14	17	21	56
Failed	1	2	8	4	15
Total	29	45	62	34	170
B. Percentage by Worst Rating					
1 or 2	26.09	26.09	39.13	8.70	
3	22.64	32.08	35.85	9.43	
4 or 5	7.14	25.00	30.36	37.50	
Failed	6.67	13.33	53.33	26.67	
Total	17.06	26.47	36.47	20.00	
C. Percentage by REST Rating					
1 or 2	41.38	26.67	29.03	11.76	27.06
3	41.38	37.78	30.65	14.71	31.18
4 or 5	13.79	31.11	27.42	61.76	32.94
Failed	3.45	4.44	12.90	11.76	8.82

Note: All banks were California banks supervised from the FDIC's Los Angeles East, Los Angeles West, and Orange County field offices. All banks had a 1 or 2 CAMELS composite rating as of December 1988. REST ratings are based on December 1988 Call Reports. The worst rating was the worst CAMELS composite rating assigned between 1992 and 1995, and all failures occurred between 1992 and 1995.

Table 3 shows that the model is not perfect but that it does correctly identify a large percentage of problem banks and failures. Consider Type I accuracy first. Panel A indicates that 15 banks failed; 4 of the 15 had REST ratings worse than 4.5, while 8 of the 15 had REST ratings between 3.5 and 4.5. Panel B shows this is 27 percent and 53 percent of the failures, respectively. Thus, if banks with REST ratings of 3.5 or worse are targeted, the Type I accuracy for failures is 80 percent. A similar analysis shows that for problem banks, REST has a Type I accuracy of 68 percent.

The analysis of Type II accuracy also shows that REST is quite accurate. Panel C indicates that among the banks with REST ratings of 4.5–5, 62 percent became problem banks and 12 percent failed; put differently, 74 percent either failed or were in danger of failing. For REST ratings of 3.5–4.5, 28 percent were problem banks, and 13 percent failed. In other words, just over 40 percent had severe difficulties.

Several points about this backtesting should be emphasized. All the banks were rated 1 or 2 at the time of the December 1988 Call Report, and the examination ratings were given three to seven years after the Call Report. In short, REST did a reasonably good job of identifying which sound banks were most likely to encounter difficulties three to seven years later.

However, the example of Southern California was chosen precisely because real estate problems were severe there. This is a critical piece of information. The stress test identifies banks that could become problems if there were a real estate crisis. REST does not identify real estate markets that are susceptible to crisis. The backtest was successful because the Southern California market did in fact have a crisis; REST did not identify that market as one vulnerable to crisis. In the jargon of forecasting, the stress test provides conditional, not unconditional, forecasts.[25]

[25] Earlier in the same period Texas had a major crisis, which we did not use for two reasons. First, large bank-holding companies present a number of difficulties because of the connections between banks in the holding company. Second, the real estate problems in Texas began after many banks in the state had already gotten into trouble because of loans to the oil and gas industry. However, tests on the 1986 data from Texas show results similar to those presented in the text for Southern California. As of December 1986, only 34 banks had a composite CAMELS rating of 1 or 2 and a REST rating of 5. Of those 34, 13 (38 percent) failed and 13 (38 percent) became problem banks. Only 1 maintained a 1 or 2 rating until 1993. In contrast, 338 banks had a REST rating of 2, and only 12 (3 percent) failed, while 43 (13 percent) became problem banks.

Table 4

Performance of Stress Test in Atlanta

Worst Exam Rating	REST Rating 1.0–2.5	2.5–3.5	3.5–4.5	4.5–5.0	Total
A. Number of Banks					
1 or 2	2	7	7	4	20
3	3	3	4	4	14
4 or 5	0	2	6	4	12
Failed	0	0	1	0	1
Total	5	12	18	12	47
B. Percentage by Worst Rating					
1 or 2	10.00	35.00	35.00	20.00	
3	21.43	21.43	28.57	28.57	
4 or 5	0.00	16.67	50.00	33.33	
Failed	0.00	0.00	100.00	0.00	
Total	10.64	25.53	38.30	25.53	
C. Percentage by REST Rating					
1 or 2	40.00	58.33	38.89	33.33	42.55
3	60.00	25.00	22.22	33.33	29.79
4 or 5	0.00	16.67	33.33	33.33	25.53
Failed	0.00	0.00	5.56	0.00	2.13

Note: All banks were headquartered in the Atlanta MSA. All banks had a 1 or 2 CAMELS composite rating as of December 1987. REST ratings are based on December 1987 Call Reports. The worst rating was the worst CAMELS composite rating assigned between 1991 and 1994, and the failure occurred between 1991 and 1994.

Both New England and Southern California suffered from extremely bad real estate problems. REST has also been backtested on episodes of less severe real estate problems. For example, table 4 reports the results (based on December 1987 data and examination ratings from the period 1991–1994) for banks headquartered in the Atlanta MSA.

The problems in Atlanta were clearly less severe than those in New England or Southern California. Only one bank failed, and no banks received a CAMELS 5 rating. Nonetheless, institutions identified by the stress test were more likely to have severe difficulties. Only 2 of the 17 institutions (12 percent) with REST ratings better than 3.5 received a CAMELS 4 rating, but 11 of the 30 institutions (37 percent) with REST ratings worse than 3.5 later became problem banks or failed. Again, all these banks were CAMELS rated 1 or 2 at year-end 1987.[26]

Forecasts Based on December 2002 Data

The stress test has been run at the FDIC since 1999, and the ratings are distributed every quarter to FDIC examiners and analysts as well as to the other banking regulatory agencies. Tables 5, 6, and 7 summarize a recent set of ratings—those based on the December 31, 2002, Call Report data. In contrast to the backtests, these tables report on all institutions regardless of CAMELS rating.[27] However, institutions with equity-to-asset ratios exceeding 30 percent and loan-to-asset ratios less than 25 percent are omitted.

[26] A handful of other backtests have been done and have produced similar results.
[27] There is a second difference as well: thrifts are included in the December 2002 data.

Table 5 reports the results by FDIC region, table 6 by state (omitting U.S. territories), and table 7 by selected MSAs.

Table 5 shows that the banks in the San Francisco and Atlanta regions are unusually vulnerable to real estate problems. Of the 699 institutions in the San Francisco region, 162 (23.2 percent) had ratings of 3.5–4.5, and 250 (35.8 percent) fell into the worst category, with ratings of 4.5–5.0. This last number is the most significant, since these are the institutions that the model identifies as especially vulnerable to real estate problems. In the Atlanta region, 244 (21.2 percent) had ratings between 3.5 and 4.5, and 332 (28.8 percent) had ratings worse than 4.5. In the rest of the nation, only 12.7 percent were rated between 3.5 and 4.5, and only 9.5 percent were rated worse than 4.5. In short, the model indicates that institutions in the West and Southeast are approximately three times more likely to be vulnerable to a real estate crisis than institutions in other parts of the country.

Table 5

REST Ratings by FDIC Region
(Based on December 2002 Call Report)

FDIC Region	1.0–1.5	1.5–2.5	2.5–3.5	3.5–4.5	4.5–5.0	Total
			A. Number			
Boston	2	174	128	34	18	356
New York	14	364	231	84	58	751
Atlanta	24	261	291	244	332	1,152
Memphis	10	217	220	118	79	644
Chicago	59	927	528	284	182	1,980
Kansas City	136	1,238	397	202	138	2,111
Dallas	37	486	252	165	189	1,129
San Francisco	17	129	141	162	250	699
Total	299	3,796	2,188	1,293	1,246	8,822
Excluding Atlanta & San Francisco	258	3,406	1,756	887	664	6,971
			B. Percentage			
Boston	0.6	48.9	36.0	9.6	5.1	
New York	1.9	48.5	30.8	11.2	7.7	
Atlanta	2.1	22.7	25.3	21.2	28.8	
Memphis	1.6	33.7	34.2	18.3	12.3	
Chicago	3.0	46.8	26.7	14.3	9.2	
Kansas City	6.4	58.6	18.8	9.6	6.5	
Dallas	3.3	43.0	22.3	14.6	16.7	
San Francisco	2.4	18.5	20.2	23.2	35.8	
Total	3.4	43.0	24.8	14.7	14.1	
Excluding Atlanta & San Francisco	3.7	48.9	25.2	12.7	9.5	

Table 5 also indicates some regions of secondary concern, notably the Dallas and Memphis regions.

In table 6 (the results reported by state) the states are ranked by the percentage of institutions with stress-test ratings worse than 4.5.[28] This table clearly indicates that the vulnerable institutions are concentrated geographically, with 6 of the top 10 states being in the San Francisco region. In addition, there are only 11 states in which 30 percent or more of the institutions are extremely vulnerable, and only 4 more in which the percentage is between 20 percent and 30 percent.

Table 7 presents the data by MSA, though it includes only MSAs where at least 10 banks or thrifts are headquartered.[29] Again, the MSAs are ranked by the percentage of institutions with stress-test ratings worse than 4.5. Only the top 20 MSAs are reported in the table, and the table confirms that these MSAs are unusual. On average, REST assigns almost 60 percent of the banks and thrifts in these MSAs a rating of 4.5 or worse, whereas for all other MSAs the comparable number is approximately 20 percent. Clearly, the FDIC should be especially concerned about the health of real estate markets in these MSAs.[30]

[28] The totals in table 5 include banks and thrifts in U.S. territories.
[29] Unfortunately, some cities with very high percentages of poor REST ratings (for example, Provo, Utah, and Fort Collins, Colo.) are excluded from the table because too few institutions are headquartered in them.
[30] Some preliminary work also shows that new banks have unusually poor REST ratings. As a group, banks that are less than three years old have REST ratings comparable to those in the MSAs listed in table 7.

Table 6

REST Ratings by State
(Based on December 2002 Call Report)

	Number of Institutions						Percentage
State	1.0–1.5	1.5–2.5	2.5–3.5	3.5–4.5	4.5–5.0	Total	4.5–5.0
AZ	1	1	5	6	23	36	63.9
NV	6	3	2	3	16	30	53.3
WA	2	14	11	18	49	94	52.1
UT	3	10	8	3	25	49	51.0
NC	0	16	15	22	48	101	47.5
OR	0	3	9	7	15	34	44.1
GA	6	64	64	56	119	309	38.5
CA	1	32	62	90	103	288	35.8
CO	4	46	21	25	52	148	35.1
FL	8	38	59	79	94	278	33.8
ID	0	2	5	6	6	19	31.6
AK	0	0	4	1	2	7	28.6
MI	1	43	49	34	41	168	24.4
SC	2	20	29	26	20	97	20.6
DC	0	2	0	2	1	5	20.0
TN	3	56	56	48	40	203	19.7
NM	1	24	11	9	11	56	19.6
DE	2	7	8	5	5	27	18.5
VA	3	24	57	31	25	140	17.9
AL	4	53	48	24	26	155	16.8
MD	6	49	35	15	19	124	15.3
TX	19	277	157	105	98	656	14.9
MO	11	169	89	60	43	372	11.6
LA	1	71	54	19	18	163	11.0
OK	13	139	63	26	28	269	10.4
WI	9	113	101	54	31	308	10.1
NH	0	14	13	0	3	30	10.0
MT	2	35	20	17	8	82	9.8
NJ	2	63	33	18	12	128	9.4
IL	29	414	149	99	69	760	9.1
KS	30	219	54	32	33	368	9.0
AR	3	62	66	29	15	175	8.6
IN	6	86	69	28	16	205	7.8
MN	21	256	111	52	37	477	7.8
KY	5	110	71	38	16	240	6.7
WY	2	23	10	9	3	47	6.4
MA	1	103	62	23	12	201	6.0
NY	0	92	66	17	11	186	5.9
MS	3	28	44	22	6	103	5.8
CT	0	23	24	6	3	56	5.4
NE	32	171	39	18	11	271	4.1
SD	8	60	13	8	3	92	3.3
OH	9	161	89	31	9	299	3.0
IA	14	299	75	27	11	426	2.6
PA	4	151	86	24	4	269	1.5
HI	0	2	5	1	0	8	0.0
ME	0	18	16	4	0	38	0.0
ND	20	64	16	5	0	105	0.0
RI	1	4	6	0	0	11	0.0
VT	0	12	7	1	0	20	0.0
Total	299	3,796	2,188	1,293	1,246	8,822	14.1

Table 7

REST Ratings by MSA
(Based on December 2002 Call Report)

			Number of Institutions		Percentage
MSA	State	Total	3.5–4.5	4.5–5.0	4.5–5.0
Atlanta	GA	76	5	64	84.2
Raleigh	NC	11	1	8	72.7
Seattle	WA	36	6	26	72.2
Grand Rapids	MI	20	5	14	70.0
Portland	OR-WA	13	2	9	69.2
Naples	FL	10	4	6	60.0
Sacramento	CA	10	1	6	60.0
Phoenix	AZ	27	4	16	59.3
Nashville	TN	21	6	12	57.1
Las Vegas	NV-AZ	24	4	13	54.2
Birmingham	AL	19	4	10	52.6
Norfolk	VA-NC	14	4	7	50.0
San Jose	CA	10	3	5	50.0
Riverside	CA	19	5	9	47.4
Dallas	TX	73	12	34	46.6
Orlando	FL	24	4	11	45.8
Stockton	CA	11	4	5	45.5
Memphis	TN-AR-MS	25	6	11	44.0
Salt Lake City	UT	32	1	14	43.8
Denver	CO	33	5	14	42.4
Top 20 MSAs		508	86	294	57.9
All MSAs		4,243	873	1,009	23.8
All but the top 20 MSAs		3,735	787	715	19.1

Note: This table includes only MSAs in which at least 10 banks or thrifts are headquartered. Data are reported for only the top 20 such MSAs.

Analysis of the December 2002 Forecasts

The results of the stress test can be analyzed much as the SCOR model is. With the SCOR model, one can attribute the reasons for a forecasted CAMELS downgrade to specific variables by comparing the bank's ratios with the median ratios of all banks currently rated 2. The same technique can be used with REST.[31]

For purposes of examining the REST ratings, we defined the benchmark as the median ratios of all institutions currently rated 1 or 2. This standard of comparison cannot be identified with any existing institution; it is a composite—the "average" institution with a 1 or 2 rating.[32] This benchmark is used to calculate "weights" that trace the reason for poor ratings back to specific ratios. The weights are in terms of percentages so they necessarily sum to 100 percent. The percentages can be negative if the ratio is better than the standard. Importantly, the weights are not used in the estimation; they are merely a method of comparing an institution that has received a poor rating with an average institution.

Tables 8 and 9 illustrate how the weighting procedure can be used to analyze a result. Each table is for a hypothetical institution. The institution described in table 8 has a stress-test rating of 4.86 but a CAMELS rating of 2 and a SCOR rating of 1.51. However, almost 12 percent of the institution's assets are construction loans, and those loans make up about 81 percent of the difference between this institution and the typical 1- or 2-rated bank. Other factors contributing to the poor REST rating are nonresidential real estate (18.52 percent of the portfolio, with a weight of 6.38 percent), multifamily housing loans (weight 5.47 percent), and C&I loans (weight 4.71 percent). This institution does have some strong points, though they are not important enough to change the stress-test rating. It holds 0.89 percent of its assets in its loan-loss reserves. These reserves have a weight of –0.64 percent, indicating that although they are a positive factor, they are

[31] See appendix 2 in Collier et al. (2003) for an explanation of the method for deriving SCOR weights. The method used by REST is slightly more complicated because some variables (for example, nonaccruing loans) can never be less than zero.

[32] SCOR uses the median ratios of the banks that received a rating of 2 within the previous year.

Table 8

Sample Stress-Test Rating, Hypothetical Bank A

Cert			Charter		
Found			State		
Stress	**4.86**		Region		
			Field Office		
CAMELS	2		SCOR	1.51	
Current Data		**Weight**			
Assets	100,000	1.87			
Growth	25.42	1.66			
SCOR Ratios			**Portfolio**	**Ratio**	**Weight**
Equity	10.20	0.94	Construction	11.97	81.02
Loss Reserves	0.89	–0.64	Nonresidential Real Estate	18.52	6.38
Loans Past Due 30–89 Days	0.62	0.23	Multifamily	3.34	5.47
Loans Past Due 90+ Days	0.04	–0.06	1–4 Family	5.96	0.63
Nonaccrual Loans	0.26	–0.33	C&I	12.56	4.71
Other Real Estate	0.00	0.00	Credit Card	1.07	–0.64
Charge-offs	0.08	0.22	Other Consumer	5.08	–0.26
Provisions for Loss	0.18	0.11	Agricultural Operating	0.15	0.00
Pretax Income	1.63	–0.43	Agricultural Real Estate	0.07	0.11
Noncore Liabilities	14.41	–0.15	Depository	0.00	0.00
Liquid Assets	36.19	–0.91	Municipality	0.00	0.00
Loans and Long-Term Securities	65.98	0.07	Leases	0.00	0.00

negligible in comparison with the size of the construction loan portfolio.

Table 9 shows an institution that has a stress-test rating of 4.88. In contrast to the rating of the bank in table 8, this rating is not driven by construction loans. In fact, the bank illustrated in table 9 has no construction loans, and the stress test evaluates this as a positive factor (weight –10.27 percent). However, the institution is concentrated in multifamily housing (weight 74.07 percent). Secondary factors include a concentration in nonresidential real estate (weight 21.33 percent) and a reliance on noncore liabilities (weight 16.74 percent). This institution is relatively large (assets of $500 million), and on balance its size is a slight negative factor (weight 7.60 percent).

Table 10 presents an overview of the weights for banks that are currently rated CAMELS 1 or 2 but have REST ratings of 4.5 or worse. The variables are ordered by the median weight. Construction loans, with a median weight of almost 75 percent, are clearly the most important factor in the model. Of the 800 institutions with ratings of 4.5 or worse, 777 have weights for construction loans that exceed 5 percent. In 16 cases (as for the hypothetical bank in table 9), construction loans are a significant positive factor and have weights that exceed –5 percent. The median bank that is identified as extremely vulnerable holds 13.05 percent of its assets as construction loans, compared with 0.50 percent of the banks that receive REST ratings of between 1.50 and 2.50.

In some cases nonresidential real estate loans, C&I loans, and multifamily housing loans are also significant risk factors. In addition, large weights are regularly assigned to low levels of liquid assets, high levels of noncore liabilities, and high levels of loans past due 30–89 days. Moreover, banks with poor ratings tend to be larger and to have grown more rapidly. Most variables seldom, if ever, have significant positive or negative weights. Mortgages on 1–4 family homes generally have a positive weight, but it is never significant.

Table 11 shows that although construction loans have the most weight, they are not the only factor driving the ratings. All institutions holding construction loans exceeding 20 percent of their total assets are identified as extremely vulnerable, but 12 institutions that have no construction loans received REST ratings of 4.5 or worse.

Table 9

Sample Stress-Test Rating, Hypothetical Bank B

Cert			Charter		
Found			State		
Stress	**4.88**		Region		
			Field Office		
CAMELS	2		SCOR	1.51	
Current Data		**Weight**			
Assets	500,000	7.60			
Growth	40.37	3.69			
SCOR Ratios			**Portfolio**	**Ratio**	**Weight**
Equity	7.61	–1.65	Construction	0.00	–10.27
Loss Reserves	0.94	–0.90	Nonresidential Real Estate	43.70	21.33
Loans past due 30–89 Days	0.10	–6.55	Multifamily	43.32	74.07
Loans past due 90+ Days	0.00	–0.16	1–4 Family	0.94	0.89
Nonaccrual Loans	0.05	0.35	C&I	0.00	–8.66
Other Real Estate	0.00	0.00	Credit Card	0.00	0.04
Charge-offs	0.02	0.91	Other Consumer	0.00	–1.35
Provisions for Loss	0.24	0.23	Agricultural. Operating	0.00	0.00
Pretax Income	2.24	–2.61	Agricultural Real Estate	0.00	0.11
Noncore Liabilities	37.65	16.74	Depository	0.00	0.00
Liquid Assets	12.27	6.35	Municipality	0.00	0.00
Loans and Long-Term Securities	87.33	–0.16	Leases	0.00	0.00

Table 10

Reasons for Ratings of 4.5–5.0

	Median Weight	Negative Factors (Weights of 5% or More)		Positive Factors (Weights of –5% or Less)		Medians	
		Number	Percent	Number	Percent	4.5–5.0	1.5–2.5
Number						800	3,565
Percentage with CAMELS 1 Rating						23.5	53.0
SCOR Rating						1.77	1.56
REST Rating						4.94	2.06
Construction Loans	73.46	777	97.1	16	2.0	13.05	0.50
Nonresidential Real Estate Loans	4.37	357	44.6	6	0.8	17.00	5.17
C&I Loans	4.34	378	47.3	38	4.8	12.60	6.11
Noncore Liabilities	3.75	339	42.4	22	2.8	23.72	13.20
Growth	2.88	248	31.0	4	0.5	32.08	7.15
Assets	2.03	210	26.3	28	3.5	162,870	62,404
Liquid Assets	1.71	55	6.9	3	0.4	21.39	36.82
Multifamily Housing Loans	0.78	104	13.0	0	0.0	0.93	0.07
Loans and Long-Term Securities	0.20	0	0.0	0	0.0	79.30	68.57
Provisions for Loan Loss	0.20	4	0.5	0	0.0	0.25	0.08
1–4 Family Mortgages	0.13	0	0.0	0	0.0	14.20	17.98
Agricultural Real Estate Loans	0.12	0	0.0	0	0.0	0.06	2.99
Charge-offs	0.08	6	0.8	21	2.6	0.08	0.07
Loans Past Due 30–89 Days	0.02	259	32.4	182	22.8	0.70	0.63
Other Real Estate	0.00	0	0.0	45	5.6	0.00	0.00
Agricultural Loans	0.00	0	0.0	0	0.0	0.00	3.28
Loans to Depositories	0.00	0	0.0	0	0.0	0.00	0.00
Loans to Municipalities	0.00	0	0.0	12	1.5	0.00	0.00
Leases	0.00	0	0.0	35	4.4	0.00	0.00
Credit Card Loans	–0.02	0	0.0	4	0.5	0.09	0.00
Loans Past Due 90+ Days	–0.04	6	0.8	0	0.0	0.01	0.04
Pretax Income	–0.12	34	4.3	42	5.3	1.46	1.47
Nonaccrual Loans	–0.13	0	0.0	42	5.3	0.19	0.11
Equity	–0.29	20	2.5	0	0.0	8.56	10.28
Other Consumer Loans	–0.32	2	0.3	0	0.0	3.51	6.24
Loan-Loss Reserves	–0.56	0	0.0	26	3.3	0.88	0.73

Table 11

REST Ratings and Construction Loans

Construction Loans as a Percentage of Assets	REST Rating					
	1.0–1.5	1.5–2.5	2.5–3.5	3.5–5.5	4.5–5.0	Total
0	203	897	123	35	12	1,270
0–5	134	2,657	2,084	548	31	5,454
5–10	2	11	239	745	159	1,156
10–15			2	90	306	398
15–20				3	164	167
20–25					74	74
25–30					21	21
30–35					20	20
35–40					6	6
40–45					3	3
45–50					2	2
60–65					1	1
65–70					1	1
Total	339	3,565	2,448	1,421	800	8,573

Table 11 also shows the reason that using the ratio of construction loans to total assets by itself is inadequate. A bank could have 7 percent of its assets in construction loans and receive almost any REST rating. If the bank has no other risk factors, it will receive a rating of 1–1.5, but if other risk factors are present, it may receive a rating of 5. Before assigning ratings, the stress test considers several aspects of a bank's operations, allowing for both mitigating and exacerbating factors. A single ratio is only one number and is meaningful only after it has been put in a broader context.[33]

Trends in Stress-Test Ratings

Figures 1 and 2 show the history of stress-test ratings since December 1986 for the United States as well as some individual states. Both figures show the percentage of institutions receiving ratings of 3.5 or worse as a percentage of all institutions with REST ratings.[34] Figure 1 shows ratings in the United States and in two states that have already been discussed—Massachusetts and California. Figure 2 shows ratings in Arizona, Georgia, and Illinois. Both figures also show a definite trend in stress-test ratings: since 1993, the ratings for the United States and for all five states have become worse.

In figure 1 the effects of the real estate crises in Massachusetts and California are clear. Large percentages of the financial institutions in both states were vulnerable in the late 1980s, and the percentages of vulnerable institutions then declined dramatically. Figure 1 also shows that institutions in the two states have followed quite different paths in the last decade. Whereas the REST ratings for California banks and thrifts have again become substantially worse than those for the United States as a whole, ratings for Massachusetts banks have generally become better.

[33] Gilbert, Meyer, and Vaughan (1999) make this point forcefully.

[34] REST uses the SCOR model to assign ratings that are comparable to CAMELS ratings. Using the data on the characteristics of banks assigned CAMELS 5 ratings after actual examinations, SCOR estimates coefficients that describe the characteristics of a 5-rated bank. In 1998, there were few banks with CAMELS 5 ratings, so for that year the SCOR characterization of a 5-rated bank relies on very little data and is consequently imprecise. This imprecision affects REST ratings worse than 4 because a rating midway between 4 and 5 draws on the characterizations of both 4-rated and 5-rated banks. The imprecision in SCOR (and REST) resulted in better ratings for banks with very poor financials. If one takes a set of very poor financial ratios and assigns a rating based on pre-1997 coefficients or coefficients estimated on data from 1999 or later, the ratings would all be similar. However, the 1998 coefficients produce better ratings for the weakest financial ratios (that is, those ratios that would have been assigned a rating worse than 4 by coefficients from other periods). The data for the worst ratings are misleading in 1998 because the coefficients for 1998 are imprecise, and the ratings based on those coefficients do not reflect the innate weakness of the banks in the worst condition.

Figure 1

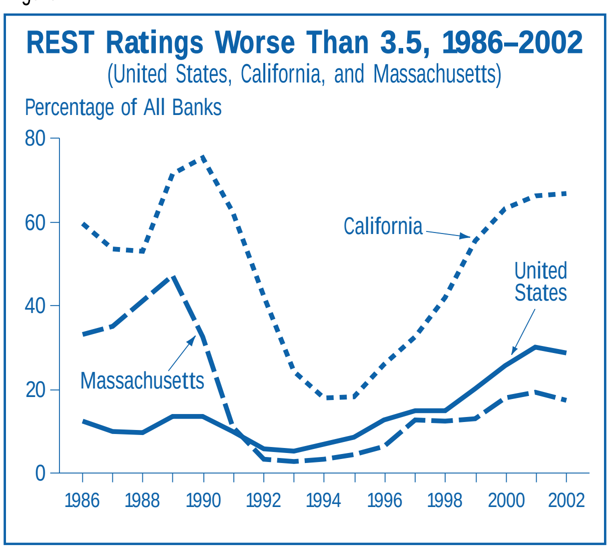


Figure 2

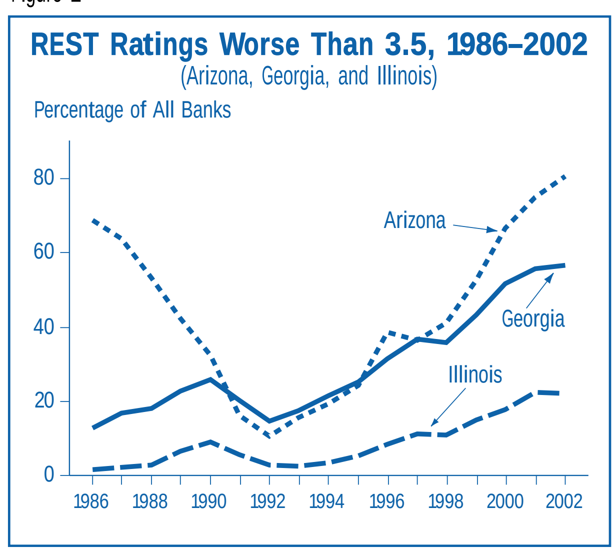


Figure 2 shows that Arizona banks and thrifts have followed a pattern similar to California's, with very poor ratings in the mid-1980s, a very rapid improvement, and a subsequent deterioration. Ratings in Georgia, in contrast, have gradually deteriorated up to the present. Georgia today has a very high percentage of banks and thrifts with poor ratings. Ratings in Illinois have followed the national pattern quite closely, with some increase before the recession of the early 1990s, a decline during the recession, and a gradual but definite increase in the percentage of poor ratings after 1993. However, ratings in Illinois have generally been a little better than ratings in the rest of the country. Both figures illustrate quite clearly that although national trends may be significant, each state has a story of its own.

Conclusion

This article has explained the development of a real estate stress test and the test's most significant results. The stress test highlights institutions whose lending practices deserve scrutiny; it therefore spotlights markets that should be inspected for evidence of incipient real estate problems. REST indicates that a large fraction of banks and thrifts in the West and the Southeast may be vulnerable to problems in the real estate market, mostly because of large concentrations in construction and development lending. REST does not, however, show that any real estate market is either overbuilt or on the verge of a crisis. There are, after all, a multitude of ways for institutions to manage and mitigate the risk of construction lending.

This article raises the questions of whether institutions that have exposures to the real estate market have adequately protected themselves and whether the real estate markets in the West and Southeast are inherently healthy. The history of banking suggests that these questions are vitally important to the FDIC.

REFERENCES

Collier, Charles, Sean Forbush, Daniel A. Nuxoll, and John O'Keefe. 2003. The SCOR System of Off-Site Monitoring: Its Objectives, Functioning, and Performance. *FDIC Banking Review* 15, no. 3:17–32.

Federal Deposit Insurance Corporation (FDIC). 1997. *History of the Eighties—Lessons for the Future*. Vol. 1, *An Examination of the Banking Crises of the 1980s and Early 1990s*. FDIC.

Gilbert, R. Alton, Andrew P. Meyer, and Mark D. Vaughan. 1999. The Role of Supervisory Screens and Econometric Models in Off-Site Surveillance. Federal Reserve Bank of St. Louis *Review* (November–December): 31–56.

Herring, Richard J., and Susan M. Wachter. 1999. Real Estate Booms and Banking Busts—An International Perspective. Occasional Paper No. 58. Group of Thirty.

Mayer, Martin. 1990. *The Greatest-Ever Bank Robbery*. Charles Scribner's Sons.