# Problem Loans and Cost Efficiency in Commercial Banks

Allen N. Berger
Board of Governors of the Federal Reserve System
Washington, DC 20551
Wharton Financial Institutions Center
Philadelphia, PA 19104 U.S.A.
Phone (202) 452-2903, Fax (202) 452-5295
Email m1anb00@frb.gov

Robert DeYoung
Office of the Comptroller of the Currency
Washington, DC 20219
Phone (202) 874-5427, Fax (202) 874-5394
Email robert.deyoung@occ.treas.gov

Forthcoming, *Journal of Banking and Finance*, Vol. 21, 1997

**Abstract:** This paper addresses a little examined intersection between the problem loan literature and the bank efficiency literature. We employ Granger-causality techniques to test four hypotheses regarding the relationships among loan quality, cost efficiency, and bank capital. The data suggest that problem loans precede reductions in measured cost efficiency; that measured cost efficiency precedes reductions in problem loans; and that reductions in capital at thinly capitalized banks precede increases in problem loans. Hence, cost efficiency may be an important indicator of future problem loans and problem banks. Our results are ambiguous concerning whether or not researchers should control for problem loans in efficiency estimation.

JEL Code: G21, G28

Key Words: Commercial Banks, Cost Efficiency, Loan Quality, Granger-causality.

---

The authors thank Bob Avery, Paul Bauer, George Benston, James Burgess, Sigbjørn Atle Berg, Gary Ferrier, Stephen Kane, Kevin Jacques, Julapa Jagtiani, Bill Lang, Tom Lutton, Knox Lovell, Joe Mason, Andrea Resti, David Runkle, Bernie Shull, Gary Whalen, and two anonymous referees for helpful comments, and Maria Montano and Joe Scalise for exceptional research support. The views expressed are those of the authors, and do not necessarily reflect those of the Board of Governors, the Office of the Comptroller of the Currency, the Department of the Treasury, or their staffs.

## I. Introduction

Over the last several years, two strands of research in the field of financial institutions have received great amounts of attention. One strand investigates the issue of problem loans. Virtually all research on the causes of bank and thrift failures find that failing institutions have large proportions of nonperforming loans prior to failure, and that asset quality is a statistically significant predictor of insolvency (see Demirguc-Kunt 1989, Whalen 1991, and Barr and Siems 1994). The other strand of research investigates the productive efficiency of financial institutions. These studies almost always find that the average institution incurs high costs and generates low profits relative to institutions on the 'best-practice' efficient frontier. Various studies of mergers, agency problems, corporate governance, branching strategies, foreign ownership, etc. offer support for a number of explanations of this inefficiency (see Berger, Hunter, and Timme 1993).

On the surface, these two topics might appear to be largely unrelated, because operations personnel typically do not participate in screening and monitoring loan customers, and because loan officers and review personnel typically do not participate in overseeing operations costs. Despite this apparent dichotomy, issues of problem loans and cost efficiency are in fact related in several important ways.

First, a number of researchers have found that failing banks tend to be located far from the best-practice frontier (e.g., Berger and Humphrey 1992, Barr and Siems 1994, DeYoung and Whalen 1994, Wheelock and Wilson 1994). Thus, in addition to having high ratios of problem loans, banks approaching failure also tend to have low cost efficiency. A number of other studies have found negative relationships between efficiency and problem loans even among banks that do not fail (Kwan and Eisenbeis 1994, Hughes and Moon 1995, Resti 1995). Cost-inefficient banks may tend to have loan performance problems for a number of reasons. For example, banks with poor senior management may have problems in monitoring both their costs **and** their loan customers, with the losses of capital generated by both these phenomena potentially leading to failure. Alternatively, loan quality problems may be caused by events exogenous to the bank, such as regional economic downturns, in which case extra expenses associated with the nonperforming loans (e.g., monitoring, negotiating workout arrangements, seizing and disposing of collateral, diverted senior managerial focus) can create the appearance, if not the reality, of low cost

efficiency.

A second empirical link between problem loans and productive efficiency appears in studies that use supervisory examination data. Peristiani (1996) and DeYoung (1997) both found measured cost efficiency to be positively related to examiners' ratings of bank management quality. Moreover, the latter study found that banks' management ratings were more strongly related to their asset quality ratings than to any of their other examination ratings. A relationship between asset quality and cost efficiency (via management quality) is consistent with the failed bank data cited above, and suggests that the negative relationship between problem loans and cost efficiency holds for the population of banks as well as for the subset of failing banks.

Third, some recent studies of bank efficiency have directly included measures of nonperforming loans in cost or production relationships. The stated purposes of this adjustment are to control for the extra costs associated with nonperforming loans and/or to control for underwriting and monitoring expenditures that influence loan quality. Berg, Førsund, and Jansen (1992) made the original observation and included nonperforming loans in a nonparametric study of bank production. Hughes and Mester (1993) and others have applied the concept to parametric estimations of bank cost functions and efficiency. As discussed below, whether or not this procedure improves the estimation of cost efficiency depends upon the underlying reason for the relationship between costs and nonperforming loans.

Thus, a number of important policy and research issues -- discovering the primary cause of problem loans and bank failures, determining the most important supervisory focus for promoting bank safety and soundness, and deciding how to estimate the cost efficiency of financial institutions -- rest on identifying the underlying relationship between problem loans and measured cost efficiency. We attempt to shed some light on these important questions by using Granger-causality analysis to test a set of hypotheses that describe the intertemporal relationships among problem loans, cost efficiency, and financial capital. We refer to these hypotheses with the mnemonics 'bad luck,' 'bad management,' 'skimping,' and 'moral hazard.' As we will see below, several of these hypotheses yield identical predictions for the contemporaneous relationships between problem loans and cost efficiency, but strikingly different intertemporal predictions. Granger-causality analysis can distinguish among these alternative cases in a

way that standard cross-section econometric analysis cannot.[1] We acknowledge that the intertemporal relationships that we measure are gross statistical associations that do not necessarily prove economic causation. Nevertheless, these associations may help reveal which among these alternative hypotheses is most compatible with the bank data.

By way of preview, the data suggest that the intertemporal relationships between problem loans and cost efficiency ran in both directions for U.S. commercial banks between 1985 and 1994. The data suggest that high levels of nonperforming loans Granger-cause reductions in measured cost efficiency, consistent with the hypothesis that the extra costs of administering these loans reduces measured cost efficiency ('bad luck'). The data also suggest that low levels of cost efficiency Granger-cause increases in nonperforming loans, consistent with the hypothesis that cost-inefficient managers are also poor loan portfolio managers ('bad management').

The remainder of the paper is organized as follows. Section II presents our four hypotheses and identifies their key empirical predictions. Section III describes the econometric model, and Section IV explains how the efficiency estimates were constructed. Section V presents the results of our Granger-causality tests, and Section VI addresses policy and research implications.

## II. Hypothesized Relationships Between Problem Loans and Measured Cost Efficiency

Each of our four hypotheses implies a different intertemporal relationship in the data. The hypotheses are not mutually exclusive, and any one of the four could dominate the behavior of a given subset of banks. The empirical analysis that follows is designed to determine which among these hypotheses may be relatively important for the banking industry as a whole. For convenience, we will describe all of the hypotheses in terms of increases in problem loans, which is the most interesting case for policy purposes.[2]

As a caveat, we note that what we actually test is whether one variable temporally precedes another variable with a particular sign or direction, conditional on lagged values of a number of variables. Our results could literally reflect 'bad luck,' 'bad management,' 'skimping' behavior, or 'moral hazard' problems, and we will refer to the empirical results in these terms for expositional convenience. However, we note that bad luck, bad management, etc. could be manifested in the data in ways other than those described

here, and that other unnamed hypotheses could also explain the empirical relationships that we find.

Under the **'bad luck'** hypothesis, external events (e.g., a local plant closing) precipitate an increase in problem loans for the bank. After the loans become past due or nonaccruing, the bank begins to expend additional managerial effort and expense dealing with these problem loans. These extra operating costs include, but are not limited to a) the additional monitoring of the delinquent borrowers and the value of their collateral, b) the expense of analyzing and negotiating possible workout arrangements, c) the costs of seizing, maintaining, and eventually disposing of collateral if default later occurs, d) the additional costs of defending the bank's safety and soundness record to bank supervisors and market participants, e) any additional precautions taken to preserve the high quality of loans that are currently performing, which becomes more crucial for a bank in a perilous financial situation, and f) the diversion of senior management attention away from solving other operations problems. Most of these costs, especially the costs associated with loan workout and default, are incurred well after the increase in problem loans. Thus, under the bad luck hypothesis, we expect increases in nonperforming loans to Granger-cause (i.e., temporally precede) decreases in measured cost efficiency. Importantly, under the bad luck hypothesis, the extra expenses associated with problem loans create the appearance, but not necessarily the reality, of lower cost efficiency. Faced with an exogenous increase in nonperforming loans, even the most cost efficient banks have to purchase the additional inputs necessary to administer these problem credits.

Under the **'bad management'** hypothesis, low measured cost efficiency is a signal of poor senior management practices, which apply to both day-to-day operations and to managing the loan portfolio. Subpar managers do not sufficiently monitor and control their operating expenses, which is reflected in low measured cost efficiency almost immediately. Managers in these banks also do not practice adequate loan underwriting, monitoring, and control. As 'bad' managers, they may a) have poor skills in credit scoring and therefore choose a relatively high proportion of loans with low or negative net present values, b) be less than fully competent in appraising the value of collateral pledged against the loans, and c) have difficulty monitoring and controlling the borrowers after loans are issued to assure that covenants are obeyed. In contrast to the almost immediate reduction in measured cost efficiency, poor underwriting and monitoring practices lead to high numbers of nonperforming loans only after some time passes, the loan

portfolio becomes seasoned, and delinquencies begin to mount.[3] Thus, under the bad management hypothesis, low cost efficiency is expected to occur before or Granger-cause higher nonperforming loans. Note that this hypothesis has the opposite temporal ordering from that predicted by the bad luck hypothesis, but both hypotheses predict that nonperforming loans will be negatively associated with cost efficiency.

Under the **'skimping'** hypothesis, the amount of resources allocated to underwriting and monitoring loans affects both loan quality and measured cost efficiency.[4] Here, the critical decision of the bank lies in the tradeoff between short-term operating costs and future loan performance problems. A bank maximizing long-run profits may rationally choose to have lower costs in the short run by skimping on the resources devoted to underwriting and monitoring loans, but bear the consequences of greater loan performance problems and the possible costs of dealing with these problems in the future. The reduced effort devoted to screening loan customers, appraising collateral, and monitoring and controlling borrowers after loans are issued makes the bank **appear** to be cost efficient in the short run because fewer operating expenses can support the same quantity of loans and other outputs. The stock of nonperforming loans remains unaffected in the short run, but as time passes, a higher proportion of borrowers become delinquent on their loans and the inattention to the loan portfolio becomes apparent. Thus, under the skimping hypothesis, the Granger-causality between measured cost efficiency and nonperforming loans has the same temporal ordering as the bad management hypothesis, but has the opposite sign -- skimping implies a **positive** Granger-causation from measured efficiency to problem loans.

The **'moral hazard'** hypothesis is the classical problem of excessive risk-taking when another party is bearing part of the risk and cannot easily charge for or prevent that risk-taking. Under this hypothesis, banks with relatively low capital respond to moral hazard incentives by increasing the riskiness of its loan portfolio, which results in higher nonperforming loans on average in the future. Thus, under the moral hazard hypothesis, we expect that low financial capital will Granger-cause high nonperforming loans. Because banks with ample capital likely do not face significant moral hazard incentives, we test the moral hazard hypothesis only for a subsample of banks with equity-to-asset ratios below the sample median. Although the moral hazard hypothesis does not describe the relationship between measured cost efficiency and problem loans, we include it for several reasons. First, moral hazard gives an alternative

explanation for nonperforming loans, so the effects of measured cost efficiency on nonperforming loans could be biased if the potential effects of capital were neglected. Second, moral hazard effects can magnify the effects of the other three hypotheses, and any of those hypotheses could be the primary cause of reduced capital and moral hazard incentives. Finally, as a leading theory of problem loans and bank failures, the moral hazard hypothesis has different policy implications than the other hypotheses.

These four hypotheses are not mutually exclusive. In an extreme case, all four hypotheses could affect the same bank at the same time. For example, bad luck could befall a poorly managed bank that also happens to be skimping on loan monitoring expenses. Any loss of capital as a result of the bad luck, bad management, and skimping might cause the bank to respond to moral hazard incentives and take increased risks. Similarly, banks responding to moral hazard incentives may take increased risks by skimping.

## III. The Econometric Model

We use Granger-causality techniques to test which of the four hypotheses are consistent with data. As with any econometric procedure, Granger-causality yields gross statistical associations that can only indicate consistency or inconsistency with an hypothesis, not proof of economic causation. However, because our hypotheses each imply unique time-ordered and signed relationships among pairs of the three important variables (nonperforming loans, cost efficiency, and capital), Granger-causality tests may indicate which of our hypotheses are generally consistent or inconsistent with the data. The Granger-causality model is specified as follows:

$$NPL_{i,t} = f_1(NPL_{i,lag}, X\text{-}EFF_{i,lag}, CAP_{i,lag}, RWA_{i,lag}, YEAR_t, REGION_i, YEAR_t \bullet REGION_i) + \epsilon_{1i,t} \quad (1)$$

$$X\text{-}EFF_{i,t} = f_2(NPL_{i,lag}, X\text{-}EFF_{i,lag}, CAP_{i,lag}, RWA_{i,lag}, YEAR_t, REGION_i, YEAR_t \bullet REGION_i) + \epsilon_{2i,t} \quad (2)$$

$$CAP_{i,t} = f_3(NPL_{i,lag}, X\text{-}EFF_{i,lag}, CAP_{i,lag}, RWA_{i,lag}, YEAR_t, REGION_i, YEAR_t \bullet REGION_i) + \epsilon_{3i,t} \quad (3)$$

The dependent variable in equation (1) is the nonperforming loan ratio $NPL_{i,t}$. This ratio equals loans that are either past due at least 90 days or in nonaccrual status divided by total loans for bank i in year t. NPL is the most commonly agreed-upon definition of problem loans in both the research literature and the trade press, and has the benefit of being subject to less managerial discretion than are other measures of loan quality, such as loan loss provisions and charge-offs.[5] The dependent variable in equation (2) is

estimated cost efficiency $X\text{-}EFF_{i,t}$, which measures the short-term X-efficiency of bank i relative to its peers in year t, or how close the bank is to the estimated industry-wide best-practice cost frontier in that year. The details of measuring X-efficiency are presented in the next section. The third dependent variable is the ratio of equity capital to assets, $CAP_{i,t}$, which measures the bank's financial cushion for absorbing loan losses. The CAP equation (3) completes the model and verifies the basic relationships that are being investigated.

The right-hand-side variables in equations (1) - (3) include lagged values of the dependent variables NPL, X-EFF, and CAP, as is standard procedure for Granger-causality models. That is, we attempt to determine whether variable $Y_i$ Granger-causes variable $Y_j$ by testing whether the past history of $Y_i$ adds information in predicting $Y_j$, after taking into account the past history of all the other Y's. A number of control variables are added to all three equations.[6] We define $RWA_{i,t}$ to be the ratio of risk-weighted assets (according to the Basle Accord) to gross total assets for bank i in year t, and include lagged values of RWA in all three equations.[7] RWA is added because certain portfolio mixes will naturally yield more nonperforming loans, because low measured cost efficiency may reflect the costs associated with a more loan-intensive balance sheet, and because capital markets and bank regulators may require banks with certain lending strategies to have higher capital ratios.[8] In addition, because NPL, X-EFF, and CAP are often strongly affected by regional or transitory conditions, we must control in at least a crude fashion for the economic environment of the bank. We therefore specify $YEAR_t$, a set of dummy variables for each year of the data (excluding a base year), which accounts for changes in the macroeconomy and in the regulatory treatment of banks over time. We also specify $REGION_i$, a set of three dummies that indicate whether the bank was located in New England, in the Southwest, or in California, because these three geographic regions suffered particularly severe economic downturns relative to the rest of the country during parts of the sample period.[9] The YEAR and REGION variables also appear interactively to account for changes in regional economic conditions over time.

The bad luck hypothesis predicts that the sum of the lagged NPL coefficients in the X-EFF equation (2) will be negative. We focus on the sum of the lagged coefficients, rather than the individual lagged coefficients, because the sum measures the total effect of past NPL on current X-EFF, and because

the sum is more accurately measured than its individual parts due to collinearity among the individual lagged values. The bad management hypothesis predicts that the sum of the coefficients on the X-EFF lags in the NPL equation (1) will be negative, but the skimping hypothesis predicts that this sum will be positive -- we view this coefficient sum as the net effect of these two opposing theories. The final hypothesis, moral hazard, predicts that the sum of the coefficients on the CAP lags in the NPL equation (1) will be negative. We test the moral hazard hypothesis only for a subsample of thinly capitalized banks, because banks with high levels of capital likely do not face significant moral hazard incentives. Changes in cost efficiency and nonperforming loans affect bank earnings, so in equation (3) we expect to find a negative sum of coefficients on the NPL lags and a positive sum of coefficients on the X-EFF lags, although we cannot distinguish among any of our four hypotheses using this equation. Such findings would extend the stylized facts from the banking literature that problem loans and inefficiencies are strongly associated with losses of capital.

We estimate equations (1) - (3) separately, applying OLS techniques to annual observations of U.S. commercial banks from 1985 through 1994. Some years have more observations than others because of failure, acquisition, or entry during the sample period. The size of the panel depends on the number of lags included in the model -- there are 69,742 observations, 57,780 observations, and 46,504 observations, respectively, when we include 3, 4, and 5 annual lags in the model. Including up to five years of lags of the dependent variables in each equation should absorb the bulk of any serial correlation in the error terms.[10] Similarly, including YEAR, REGION, and YEAR*REGION absorbs any cross-section correlation that is constant across banks within a year, across banks within a region, or across banks in a region in a particular year. Also, because the same right-hand-side variables are specified for each of the three equations, there is no need to take account of the correlations across the error terms in the different equations. In such a specification, OLS and GLS yield identical estimators.

Table 1 shows definitions, sample means, and standard deviations for the variables used in the four-lag Granger-causality model. NPL, CAP, and RWA are December-to-December averages, and are constructed using data from the annual Reports of Condition and Income (call reports). We use annual observations because the end-of-year call report data are more accurate and smooth-out short-run

aberrations in the data better than do quarterly data.[11]

## IV. Estimating Cost Efficiency

We construct the NPL and CAP ratios directly from the data, but we must estimate the X-EFF ratio. An efficient operating cost frontier is specified as follows for each year t in our analysis:

$$
\begin{aligned}
lnOC \;=\; & \alpha_0 + \sum_{j}^{5} \beta_j lnY_j + \frac{1}{2}\sum_{j}^{5}\sum_{k}^{5} \beta_{jk} lnY_j lnY_k + \sum_{m}^{2} \gamma_m lnW_m \\
& + \frac{1}{2}\sum_{m}^{2}\sum_{n}^{2} \gamma_{mn} lnW_m lnW_n + \sum_{j}^{5}\sum_{m}^{2} \rho_{jm} lnY_j lnW_m \\
& + \sum_{j=1}^{7} [\delta_j \cos Z_j + \theta_j sinZ_j] + \sum_{j=1}^{7}\sum_{k=j}^{7} [\delta_{jk} \cos(Z_j + Z_k) + \theta_{jk} \sin(Z_j + Z_k)] \\
& + \sum_{j=1}^{7}\sum_{k=j}^{7}\sum_{l=k}^{7} [\delta_{jkl} \cos(Z_j + Z_k + Z_l) + \theta_{jkl} \sin(Z_j + Z_k + Z_l)] \\
& + \lambda_L LIMIT + \lambda_U UNIT + \eta \qquad (4)
\end{aligned}
$$

We estimate ten separate annual cost frontiers, rather than estimating a common frontier across time, because the technology or bank that is most efficient in one year may not be the most efficient in another year.[12] OC is operating (noninterest) expense; Y is a vector of outputs including commercial loans, consumer loans, real estate loans, transactions deposits, and fee-based income;[13] W is a vector of input prices including the prices of labor and physical capital;[14] and LIMIT and UNIT are, respectively, dummy variables equal to 1 for banks in states that restricted or banned branch banking during the year being observed. All of the cost function variables were constructed using end-of-year call report data, except for UNIT and LIMIT which were constructed from information in Amel (1993).[15] We estimate an operating cost function (i.e., noninterest expense only) rather than a total cost function because operating costs have been shown elsewhere to comprise the bulk of cost inefficiency at banks (see Berger and Humphrey 1991); because the costs associated with credit evaluation, loan monitoring, and administering

to problem loans are mainly noninterest expenses; and because interest costs may be substantially influenced by the local market power of the bank.[16] The Granger-causality results were robust to specifying a total cost function (including interest expenses) in place of the operating cost function.

The cost equation (4) is specified as a Fourier-flexible functional form, which combines a standard translog form with the non-parametric Fourier form. Because it includes trigonometric transformations (i.e., cos*Z* and sin*Z*) of the variables, the Fourier-flexible form can globally approximate the underlying cost function over the entire range of data.[17] This is a theoretical improvement on the translog form, which is a local approximation that may perform poorly for observations far from the sample means. In application, McAllister and McManus (1993), Berger, Leusner, and Mingo (1994), and Mitchell and Onvural (1996) have all found that the Fourier-flexible form fits the data better than the translog form for banking data.

We use the stochastic cost frontier approach (SFA) to generate an annual estimate of cost efficiency for each bank.[18] The error term $\eta$ is a composite expression, $\eta = \ln U + \ln V$, where $\ln U$ captures cost inefficiency and is assumed to follow a (positive) truncated normal distribution[19], and $\ln V$ captures random error and is assumed to follow a symmetric normal distribution. We construct the cost efficiency variable X-EFF as follows. Using equation (4), we estimate the expected value of the inefficiency term $\ln U$ conditional on the estimated value of the residual $\eta$.[20] For bank i in year t, $\text{X-EFF}_{i,t}$ equals $U_{min,t}/U_{i,t}$, the estimated value of U for the most efficient bank at time t divided by the estimated value of U for bank i, where $U = e^{\ln U}$. Thus, X-EFF increases with cost efficiency and ranges between 0 and 1.

As shown in Table 1, the average bank is measured to be about 92 percent efficient over the entire sample period, a higher level of cost efficiency than is found in most other studies of commercial banks. This finding may be because the Fourier-flexible, truncated normal specification used here is more general than the standard translog, half normal model typically used in the literature -- the translog functional form is a special case of the Fourier-flexible in which the trigonometric terms are set to zero, and the half normal distribution is a special case of the truncated normal in which the corresponding normal distribution is truncated at the center. To examine whether existing studies may have understated efficiency by imposing too much structure on the cost function and the inefficiency term $\ln U$, we re-estimated X-EFF using the standard translog and half normal restrictions. This yielded statistically significantly lower estimates of X-

EFF in which cost inefficiency was about twice as large as our original estimates. This suggests that existing studies may have incorrectly attributed too much of the variability of costs to inefficiency because of excessively restrictive specifications.[21] Our Granger-causality results are materially unaffected by these specification differences -- these two sets of estimated X-EFF have a rank order correlation of .969, which suggests that most of any specification error affects the levels, but not the bank-specific ordering, of estimated efficiency.

We purposely do not include nonperforming loans in the cost function. For a given bank, nonperforming loans may be to some extent either exogenous or endogenous -- it is exogenous to the extent that the nonperformance is from bad luck, and endogenous to the extent that the nonperformance is from bad management or skimping (i.e., due to actions taken by management). To test among these hypotheses, it is important to measure efficiency without controlling for nonperforming loans. In the Granger-causality model, we need to test for the relationships between X-EFF and NPL using a measure of X-EFF that has not previously been conditioned on NPL. Moreover, the relationships between X-efficiency and bank failure, management quality, etc. in the existing literature are drawn primarily from cost functions that did not control for loan quality.

## V. Results

Table 2 displays the coefficient estimates from equations (1)-(3) using four lags on the right-hand-side variables. All of the major results are robust to including either three or five lags (not shown). In the X-EFF equation, the sum of the lagged NPL coefficients is -.0599 and is significant at the 1 percent level. This result is consistent with the notion that after loans become past due or nonaccruing, operating costs rise because of the difficulty in dealing with these loans, as predicted by the bad luck hypothesis. The economic impact of this result is typically small -- a one standard deviation increase in NPL (from .0168 to .0377) predicts a cumulative reduction in measured cost efficiency over four years from .9224 to .9211, or a 1.7 percent increase in predicted inefficiency. However, the impact may be greater for individual banks that experience larger changes in nonperforming loans or have relatively high costs of dealing with nonperforming loans. Note that the lagged coefficient sums on CAP and RWA were not significantly different from zero in this equation.[22]

The NPL equation suggests that the bad management hypothesis dominates the skimping hypothesis for the data set as a whole. The sum of the coefficients on the lagged X-EFF variables equals -.0068 and is significant at the 1 percent level. This suggests that after measured cost efficiency declines, nonperforming loans increase, possibly because of poor loan portfolio management as predicted by the bad management hypothesis. The typical economic impact of this result is also small -- a one standard deviation reduction in X-EFF (from .9224 to .8843) predicts a cumulative increase in the nonperforming loan ratio over four years from .0168 to only .0171, or a 1.8 percent increase in nonperforming loans. Again, this effect could be larger for some banks. The lagged RWA coefficient sums are positive and significant in the NPL equation, which suggests that a relatively risky loan portfolio mix will eventually yield relatively high numbers of nonperforming loans.

Although the bad management hypothesis dominates the skimping hypothesis on average for the entire sample, this does not preclude the possibility of skimping in individual banks. For example, we might expect to find concentrations of 'skimpers' among the most cost-efficient banks, i.e., banks that willingly trade loan quality for cost reductions, but manage the resulting loan quality problems in a cost effective fashion. To investigate this possibility, we re-estimated the NPL equation for the subsample of banks with X-EFF(-1), the first lagged value of X-EFF, greater than the median in **every** year of our analysis.[23] Partial results of these regressions are shown in the first panel in Table 3. In the NPL equation, the sum of the lagged X-EFF coefficients equals .0233 and is significant at the one percent level, suggesting that measured cost efficiency positively Granger-causes nonperforming loans among highly efficient banks, supporting the skimping hypothesis. For the typical bank in this subsample, a one standard deviation increase in X-EFF (from .9522 to .9675) predicts a cumulative increase in the nonperforming loan ratio over four years from .0128 to .0132, or a 3.1 percent increase in nonperforming loans.

We test the moral hazard hypothesis by estimating the NPL equation separately using only the annual observations for which CAP(-1) was below the sample median. Partial results of this regression are shown in the second panel in Table 3. The sum of the coefficients on the lagged CAP variables equals -.0510 and is significant at the 1 percent level. These data support the moral hazard hypothesis, and suggest that, on average, thinly capitalized banks take increased portfolio risk, which results in higher

levels of problem loans in the future. For the typical low-capital bank, a one standard deviation reduction in CAP (from .0712 to .0578) predicts a cumulative increase in the nonperforming loan ratio over four years from .0186 to .0193, or a 3.8 percent increase in nonperforming loans. Again, this result likely understates the reduction in loan quality for individual banks that are subject to significant moral hazard incentives.[24]

In the CAP equation, the sum of the coefficients on the lagged X-EFF variables was positive, evidence that cost-inefficient banks are likely to have low or even negative earnings which Granger-cause reductions in capital. The sum of the lagged RWA coefficients is negative and significant, suggesting that higher risk lending strategies also Granger-cause reductions in capital. However, the sum of the lagged NPL coefficients is positive and significant, an unexpected result suggesting that high levels of nonperforming loans Granger-cause high capital ratios.[25] To investigate the possibility that well capitalized and thinly capitalized banks respond differently to increases in problem loans, we re-estimated the CAP equation for subsets of banks above and below the annual sample medians for CAP(-1). Partial results of these regressions are shown in the third and fourth panels in Table 3. We continue to find that nonperforming loans positively Granger-cause capital ratios at the low-capital banks, which suggests that these banks take action to replenish capital after nonperforming loans increase, perhaps under pressure from regulators or capital markets. We find no Granger-causality between nonperforming loans and capital at the high-capital banks.

## VI. Conclusions

The results of our analysis suggest that the intertemporal relationships between loan quality and cost efficiency run in both directions. The data provide support for the bad luck hypothesis -- increases in nonperforming loans tend to be followed by decreases in measured cost efficiency, suggesting that high levels of problem loans cause banks to increase spending on monitoring, working out, and/or selling off these loans, and possibly become more diligent in administering the portion of their existing loan portfolio that is currently performing. For the industry as a whole, the data favor the bad management hypothesis over the skimping hypothesis -- decreases in measured cost efficiency are generally followed by increases in nonperforming loans, evidence that bad management practices are manifested not only in excess

expenditures, but also in subpar underwriting and monitoring practices that eventually lead to nonperforming loans. However, for a subset of banks that were consistently efficient across time, the data favor the skimping hypothesis -- increases in measured cost efficiency generally precede increases in nonperforming loans, suggesting that these banks purposely trade short-run expense reductions for long-run reductions in loan quality. Finally, decreases in bank capital ratios generally precede increases in nonperforming loans for banks with low capital ratios, evidence that thinly capitalized banks may respond to moral hazard incentives by taking increased portfolio risks. Each of these results has a relatively small effect on banks on average, but may have a substantial effect on individual banks that are most subject to bad luck, bad management, skimping, and/or moral hazard.

Depending on whether or not these results are confirmed by future research, our findings may have implications for economic policy. The bad luck hypothesis suggests that bank failures are caused primarily by uncontrollable external events, and implies that prudential regulation and supervision could reduce the risk of failure by limiting banks' exposures to external shocks (e.g. limits on loan concentrations, allowing interregional diversification through interstate mergers and loan sales, or encouraging low loan-to-asset ratios) or by better insulating banks from external shocks (e.g., requiring high levels of capital). In contrast, the bad management hypothesis implies that the major risks facing financial institutions are caused internally. This suggests that bank supervision and research should consider cost efficiency along with other traditional predictors of troubled banks such as loan losses and credit risk. The skimping hypothesis also identifies an internal source for risk, but implies that supervisors pay special attention to banks' internal credit control procedures (e.g., loan review, collateral appraisal). The moral hazard hypothesis implies that bank supervisors should monitor capital ratios carefully and require actions to raise the ratios quickly when they become low, similar to the intent of the prompt corrective action feature of the FDIC Improvement Act of 1991. Additional implications of this hypothesis suggest that bank regulators and researchers continue their traditional focus on the effects of capital, how to measure it, and how to better set and enforce capital requirements.

The empirical results also have implications for estimating the efficiency of financial institutions. Under the bad luck hypothesis, loan quality is driven by external events, and as such efficiency

measurement should control for nonperforming loans in cost and profit functions. This would help remove by statistical means the extra costs of dealing with nonperforming loans -- which were caused by bad luck, not by managerial inefficiency -- rather than erroneously counting these extra costs as inefficiency. Under the bad management and skimping hypotheses, however, loan quality is driven by internal events. As a result, controlling for nonperforming loans in cost and profit functions will artificially increase measured efficiency by removing statistically the part of the cost inefficiencies (or revenue deficiencies) that are correlated with inefficient portfolio management. Neither hypothesis clearly dominates the other. Ultimately, whether or not one controls for loan quality should rest on the particular efficiency application at hand.

The skimping hypothesis, which received much less empirical support, has two unique implications for measuring the efficiency of banks. First, since skimping is a long-run strategy, efficiency is more appropriately measured using several years of data. Years in which skimping saves on costs would then be averaged with years in which the resulting increase in nonperforming loans potentially forces up costs. Second, under the skimping hypothesis output quality is a choice variable. Because output quality affects both costs (underwriting and monitoring costs) and revenues (lost revenues from nonperforming loans), the skimping hypothesis would favor estimation of long-run profit efficiency.

As a by-product of our analysis, we find that our average X-efficiency estimates substantially exceed those found in most other studies of commercial bank efficiency. Further investigation suggested that much of this difference reflects our more general specification of the cost function (Fourier-flexible rather than the translog form) and the distribution of the inefficiency term (truncated normal distribution rather than the half normal). These results imply that prior studies using more restrictive specifications may have incorrectly attributed too much of the variation in costs to the estimates of cost inefficiency. Because the specification errors appear to affect primarily the average level but not the rank ordering of estimated efficiency, they do not affect our Granger-causality tests.

We stress the limitations of our analysis. The intertemporal relationships revealed by Granger-causality techniques are gross statistical associations only, and do not necessarily prove economic causation. However, these relationships are indicative of which among the alternative hypotheses are most

consistent with the data. Future research might use other statistical techniques to reveal the intertemporal relationships between loan quality and productive efficiency in financial institutions; attempt to decompose the determinants of loan quality into endogenous versus exogenous factors; or focus on the empirical consequences of controlling for loan quality when estimating efficiency.

**References**


Amel, Dean, 1993, "State Laws Affecting Commercial Bank Branching, Multibank Holding Company Expansion, and Interstate Banking," Board of Governors of the Federal Reserve System, working paper.

Avery, Robert B., and Allen N. Berger, 1991, "Risk-Based Capital and Deposit Insurance Reform,"*Journal of Banking and Finance* 15, 847-874.

Barr, Richard, and Thomas Siems, 1994, "Predicting Bank Failure Using DEA to Quantify Management Quality," Federal Reserve Bank of Dallas, Financial Industry Studies Working Paper No. 1-94.

Bauer, Paul W., 1990, "Recent Developments in the Econometric Estimation of Frontiers," *Journal of Econometrics* 46, 39-56.

Berg, Sigbjørn, Finn Førsund, and Eilev Jansen, 1992, "Malmquist Indices of Productivity Growth during the Deregulation of Norwegian Banking, 1980-89," *Scandinavian Journal of Economics* 94, 211-228.

Berger, Allen N., 1993, "Distribution-Free Estimates of Efficiency in the U.S. Banking Industry and Tests of the Standard Distributional Assumptions," *Journal of Productivity Analysis* 4, 261-292.

Berger, Allen N., 1995, "The Relationship between Capital and Earnings in Banking," *Journal of Money, Credit, and Banking* 27, 432-56.

Berger, Allen N., and Robert DeYoung, 1995, "Problem Loans and Cost Efficiency in Commercial Banks," Office of the Comptroller of the Currency, Economics and Policy Analysis Working Paper 95-5.

Berger, Allen N., and David B. Humphrey, 1991, "The Dominance of Inefficiencies over Scale and Product Mix Economies in Banking," *Journal of Monetary Economics* 28, 117- 148.

Berger, Allen N., and David B. Humphrey, 1992, "Measurement and Efficiency Issues in Commercial Banking," in Z. Griliches, ed., *Output Measurement in the Service Sectors*, National Bureau of Economic Research, Studies in Income and Wealth, Vol. 56, (University of Chicago Press, Chicago) 245-79.

Berger, Allen N., William Hunter, and Stephen Timme, 1993, "The Efficiency of Financial Institutions: A Review and Preview of Research Past, Present, and Future,"*Journal of Banking and Finance*, 17: 221-249.

Berger, Allen N., John H. Leusner, and John J. Mingo, 1994, "The Efficiency of Bank Branches," Board of Governors of the Federal Reserve, Finance and Economics Discussion Series Paper 94-26.

Cordell, Lawrence R., and Kathleen K. King, 1995, "A Market Evaluation of the Risk-Based Capital Standards for the U.S. Financial System," *Journal of Banking and Finance* 19, 531-562.

Demirguc-Kunt, Asli, 1989, "Deposit-Institution Failures: A Review of the Empirical Literature," Federal Reserve Bank of Cleveland, *Economic Review*, Quarter 4.

DeYoung, Robert, "X-Efficiency and Management Quality in National Banks," Office of the Comptroller of the Currency, *Journal of Financial Services Research* (forthcoming).

DeYoung, Robert, and Gary Whalen, 1994, "Is a Consolidated Banking Industry a More Efficient Banking Industry?," Office of the Comptroller of the Currency, *Quarterly Journal*, volume 13, number 3.

Federal Financial Institutions Examination Council, 1984-1993, *Reports of Condition and Income*.

Faulhaber, Gerald R., 1995, "Banking Markets: Productivity, Risk, and Customer Satisfaction," The Wharton School, University of Pennsylvania, working paper.

General Accounting Office, 1990, *Report to the Congress on the Bank Insurance Fund*, GAO/AFMD-90-100, September.

Hughes, Joseph P., and Loretta J. Mester, 1993, "A Quality and Risk-Adjusted Cost Function for Banks: Evidence on the 'Too-Big-To-Fail' Doctrine," *The Journal of Productivity Analysis* 4, 293-315.

Jones, David S., and Kathleen K. King, 1995, "The Implementation of Prompt Corrective Action," *Journal of Banking and Finance* 19, 491-510.

Keane, Michael P., and David Runkle, 1992, "On the Estimation of Panel-Data Models With Serial Correlation When Instruments Are Not Strictly Exogenous," *Journal of Business and Economic Statistics* 10, 1-9.

Kwan, Simon, and Robert Eisenbeis, 1994, "An Analysis of Inefficiencies in Banking: A Stochastic Cost Frontier Approach," working paper.

McAllister, Patrick H., and Douglas McManus, 1993, "Resolving the Scale Efficiency Puzzle in Banking," *Journal of Banking and Finance* 17, 389-405.

Mitchell, Karlyn, and Nur M. Onvural, 1996, "Economies of Scale and Scope at Large Commercial Banks: Evidence from the Fourier Flexible Functional Form," *Journal of Money, Credit, and Banking* 28(2), 178-199.

Peristiani, Stavros, 1996, "Evaluating the Postmerger X-Efficiency and Scale Efficiency of U.S. Banks," Federal Reserve Bank of New York, working paper.

Resti, Andrea, 1995, "Linear Programming and Econometric Methods for Bank Efficiency Evaluation: An Empirical Comparison Based on a Panel of Italian Banks," working paper.

Stevenson, Rodney E., 1980, "Likelihood Functions for Generalized Stochastic Frontier Estimation," *Journal of Econometrics* 13, 58-66.

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

Wheelock, David C. and Paul W. Wilson, 1995, "Explaining Bank Failures: Deposit Insurance, Regulation, and Efficiency," *Review of Economics and Statistics* 77, 689-700.

**Table 1**
**Definitions and descriptive statistics for variables in the Granger-causation model.**
(Statistics are for the four-lag model, N = 57,655.)

| | | mean | std. dev. |
|---|---|---|---|
| NPL | Nonperforming Loan Ratio = dollar value of loans that are either 90 days past-due or are no longer accruing interest, divided by the value of total loans. | 0.0168 | 0.0209 |
| X-EFF | X-Efficiency = short-term cost efficiency, i.e., percent of maximum cost efficiency achieved by bank based on the estimated best-practice cost frontier for the year in question. | 0.9224 | 0.0381 |
| CAP | equity capital ratio = total equity capital divided by gross total assets. | 0.0892 | 0.0292 |
| RWA | risk-weighted asset ratio = estimated ratio of total risk-weighted assets to gross total assets. | 0.5661 | 0.1275 |
| | | % of observations | |
| NE | = 1 if bank is located in Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, or Vermont; = 0 otherwise. | 1.49 % | |
| SW | = 1 if bank is located in Louisiana, Texas, or Oklahoma; = 0 otherwise. | 14.53 % | |
| CA | = 1 if bank is located in California; = 0 otherwise. | 2.49 % | |
| YEAR(89) | = 1 in 1989; = 0 otherwise. | 17.60 % | |
| YEAR(90) | = 1 in 1990; = 0 otherwise. | 17.33 % | |
| YEAR(91) | = 1 in 1991; = 0 otherwise. | 17.07 % | |
| YEAR(92) | = 1 in 1992; = 0 otherwise. | 16.69 % | |
| YEAR(93) | = 1 in 1993; = 0 otherwise. | 16.28 % | |
| YEAR(94) | = 1 in 1994; = 0 otherwise. | 15.03 % | |

**Table 2 -- Granger-causality tests.**
Four lags.

| | (1) NPL | (2) X-EFF | (3) CAP |
|---|---|---|---|
| INTERCEPT | .0011 | .2266** | -.0032* |
| | (0.61) | (83.92) | (2.43) |
| NPL(-1) | .5626** | -.1316** | -.0146** |
| | (140.45) | (22.99) | (5.19) |
| NPL(-2) | .0593** | .0099 | .0142** |
| | (13.56) | (1.58) | (4.61) |
| NPL(-3) | .0084* | .0351** | .0092** |
| | (2.09) | (6.13) | (3.27) |
| NPL(-4) | .0134** | .0267** | .0075** |
| | (4.18) | (5.84) | (3.33) |
| NPL(total) | .6437** | -.0599** | .0163** |
| | (183.31) | (11.94) | (6.70) |
| X-EFF(-1) | -.0164** | .5940** | .0070** |
| | (6.34) | (160.98) | (3.89) |
| X-EFF(-2) | .0056 | .1809** | -.0022 |
| | (1.83) | (41.50) | (1.06) |
| X-EFF(-3) | .0062 | -.0162** | .0045* |
| | (1.94) | (3.54) | (2.02) |
| X-EFF(-4) | -.0022 | .0032 | .0055** |
| | (0.84) | (0.86) | (3.00) |
| X-EFF(total) | -.0068** | .7619** | .0148** |
| | (3.35) | (263.04) | (10.44) |
| CAP(-1) | -.0540** | -.0538** | .9358** |
| | (9.21) | (6.41) | (227.47) |
| CAP(-2) | .0418** | .0521** | -.0225** |
| | (5.11) | (4.46) | (3.92) |
| CAP(-3) | .0305** | .0119 | .0131* |
| | (3.67) | (1.01) | (2.24) |
| CAP(-4) | .0011 | -.0056 | .0074 |
| | (0.19) | (0.65) | (1.74) |
| CAP(total) | .0194** | .0046 | .9338** |
| | (7.63) | (1.27) | (521.84) |
| RWA(-1) | .0102** | .0178** | -.0184** |
| | (8.21) | (10.05) | (21.15) |
| RWA(-2) | .0057** | -.0171** | .0063** |
| | (3.51) | (7.38) | (5.51) |
| RWA(-3) | .0016 | -.0032 | .0017 |
| | (1.02) | (1.40) | (1.53) |
| RWA(-4) | -.0035** | .0014 | .0028** |
| | (2.96) | (0.85) | (3.43) |
| RWA(total) | .0140** | -.0011 | -.0076** |
| | (23.95) | (1.29) | (18.14) |

Absolute values of t-statistics are in parentheses ( ).
** and * indicate significance at the 1 and 5 percent levels, respectively.

## Table 2 -- continued

| | (1) NPL | (2) X-EFF | (3) CAP |
|---|---|---|---|
| YEAR(90) | .0018** | -.0207** | -.0019** |
| | (6.39) | (51.41) | (9.48) |
| YEAR(91) | .0012** | -.0077** | .0007** |
| | (4.15) | (19.49) | (3.77) |
| YEAR(92) | -.0018** | -.0065** | .0020** |
| | (6.76) | (17.18) | (10.94) |
| YEAR(93) | -.0020** | .0080** | .0033** |
| | (7.28) | (20.21) | (17.05) |
| YEAR(94) | -.0024** | .0147** | -.0002 |
| | (8.38) | (36.59) | (1.23) |
| NE | .0155** | .0019 | -.0015* |
| | (13.98) | (1.20) | (1.96) |
| NE*YEAR(90) | .0049** | .0036 | -.0022 |
| | (3.06) | (1.58) | (1.90) |
| NE*YEAR(91) | -.0023 | -.0098** | -.0046** |
| | (1.36) | (4.12) | (3.94) |
| NE*YEAR(92) | -.0136** | -.0064** | .0014 |
| | (7.91) | (2.63) | (1.14) |
| NE*YEAR(93) | -.0115** | -.0016 | .0024* |
| | (6.56) | (0.63) | (1.99) |
| NE*YEAR(94) | -.0126** | .0003 | .0027* |
| | (7.02) | (0.12) | (2.12) |
| SW | .0084** | .0056** | -.0036** |
| | (19.16) | (8.99) | (11.77) |
| SW*YEAR(90) | -.0056** | -.0088** | .0009* |
| | (9.14) | (10.05) | (2.05) |
| SW*YEAR(91) | -.0063** | -.0090** | .0028** |
| | (10.36) | (10.30) | (6.50) |
| SW*YEAR(92) | -.0063** | -.0088** | .0043** |
| | (10.27) | (10.07) | (10.03) |
| SW*YEAR(93) | -.0070** | -.0044** | .0046** |
| | (11.33) | (5.04) | (10.54) |
| SW*YEAR(94) | -.0070** | -.0022* | .0029** |
| | (11.15) | (2.48) | (6.50) |
| CA | -.0023* | .0065** | .0024** |
| | (2.07) | (4.18) | (3.11) |
| CA*YEAR(90) | .0013 | .0001 | .0033** |
| | (0.83) | (0.05) | (3.07) |
| CA*YEAR(91) | .0139** | -.0098** | -.0034** |
| | (9.44) | (4.66) | (3.34) |
| CA*YEAR(92) | .0146** | -.0126** | -.0026* |
| | (10.14) | (6.10) | (2.56) |
| CA*YEAR(93) | .0157** | -.0128** | -.0053** |
| | (11.03) | (6.26) | (5.34) |
| CA*YEAR(94) | .0071** | -.0033 | .0001 |
| | (4.96) | (1.63) | (0.11) |
| Adjusted $R^2$ | .4944 | .6892 | .8722 |
| N | 57,655 | 57,655 | 57,655 |

Absolute values of t-statistics are in parentheses ( ).
** and * indicate significance at the 1 and 5 percent levels, respectively.

**Table 3 -- Granger-causality tests.**
Partial Results for Subsample Estimations. Four lags.

| | coefficient estimates | t-statistics (absolute values) |
|---|---|---|
| Dependent variable is NPL. | | |
| Banks with X-EFF(-1) greater than sample median in every year (N = 12,756, adjusted $R^2$ = .4943). | | |
| X-EFF(-1) | .0002 | (0.02) |
| X-EFF(-2) | .0117 | (1.00) |
| X-EFF(-3) | .0133 | (1.17) |
| X-EFF(-4) | -.0019 | (0.25) |
| X-EFF(total) | .0233** | (3.13) |
| Dependent variable is NPL. | | |
| Banks with CAP(-1) less than sample median in individual years (N = 28,826, adjusted $R^2$ = .5103). | | |
| CAP(-1) | -.1575** | (11.99) |
| CAP(-2) | .0855** | (6.20) |
| CAP(-3) | .0084 | (0.65) |
| CAP(-4) | .0126 | (1.35) |
| CAP(total) | -.0510** | (4.86) |
| Dependent variable is CAP. | | |
| Banks with CAP(-1) less than sample median in individual years (N = 28,826, adjusted $R^2$ = .5039). | | |
| NPL(-1) | -.0302** | (8.96) |
| NPL(-2) | .0160** | (4.32) |
| NPL(-3) | .0110** | (3.20) |
| NPL(-4) | .0112** | (4.07) |
| NPL(total) | .0080* | (2.58) |
| Dependent variable is CAP. | | |
| Banks with CAP(-1) greater than sample median in individual years (N = 28,826, adjusted $R^2$ = .8557). | | |
| NPL(-1) | -.0068 | (1.45) |
| NPL(-2) | .0067 | (1.34) |
| NPL(-3) | .0024 | (0.53) |
| NPL(-4) | -.0002 | (0.05) |
| NPL(total) | .0021 | (0.52) |

** and * indicate significance at the 1 and 5 percent levels, respectively.

**Footnotes**

1. Although application of Granger-causality methods to microeconomic banking data is unusual, it is not unprecedented. See Berger (1995).

2. See our longer working paper (Berger and DeYoung 1995) for more details about the intertemporal patterns and policy and research implications of these four hypotheses.

3. This delay will be exacerbated if managers do not immediately record nonperforming loans. The General Accounting Office (1990) found evidence that troubled banks systematically underreport nonperforming loans.

4. Berg, Førsund, and Jansen (1992), Hughes and Mester (1993), and others use nonperforming loan to proxy for loan quality, chiefly because other dimensions of service and product quality are difficult to measure. Faulhaber (1995) appeals to survey data, and shows that increases in customer satisfaction are associated with slightly increased costs.

5. NPL is difficult to manipulate by mangers because loans must be reported as nonperforming if the borrower is past due at least 90 days. Nevertheless, manipulation of NPL could take place by lending the borrower additional funds to make payments so that the loan does not become past due 90 days. To the extent that banks delay reporting nonperforming loans, the effects may still be captured in some of our lagged coefficients.

6. Had we not included these control variables, our model would be equivalent to a vector-autoregressive (VAR) specification in three variables.

7. RWA is designed to capture credit risk taking into account a number of factors such as the type of borrower and the existence of collateral and guarantees. The risk weights have been shown to be positively related to the probability of failure and to some accounting and market measures of risk, although the correspondence has been shown to be relatively weak (see Avery and Berger 1991, Cordell and King 1995, Jones and King 1995).

8. Including lagged values of RWA generally does not control for moral hazard behavior. When a bank

responds to moral hazard incentives by taking increased risks, these risks are likely taken in secret by substituting higher risk loans for lower risk loans **within** a risk-weighted asset loan category, which is not reflected in RWA.

9. New England includes Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont. The Southwest includes Louisiana, Oklahoma, and Texas.

10. We acknowledge the possibility of bias if not all the serial correlation is absorbed by the lags. See, for example, Keane and Runkle (1992).

11. By using annual data, however, we may understate the magnitude of any Granger causation among NPL, X-EFF, and CAP if a substantial amount of the intertemporal effects associated with our four hypotheses occur within a year.

12. We estimated equation (4) using maximum likelihood techniques, and imposed the standard symmetry and homogeneity restrictions on the translog portion of the model. Factor share equations were omitted because application of the usual cross-equation restrictions would impose the assumption that the given input proportions were the allocatively efficient ones (see Berger 1993, p. 266).

13. Transactions deposits include demand deposits, NOW accounts, automatic transfer service accounts, and telephone and pre-authorized transfer accounts. Fee-based income equals gross noninterest income less both service charges on deposit accounts and gains(losses) from securities and foreign exchange trading. Fee-based income is included to control for activities other than deposits and loans such as off-balance sheet activities and trust services.

14. The price of labor equals salaries and benefits divided by the number of full time equivalent workers. The price of physical capital equals expenditures on equipment and premises divided by the book value of physical assets.

15. Operating costs (OC) and the numerators used to construct the input prices (W) are flow variables that reflect accumulated activity over the course of each year, while the output vector (Y) and the input price denominators are averages of beginning-of-year and end-of-year values. Because the natural log of zero is

undefined, we added a small positive amount (1, which represents $1,000 of output) to each of the elements of Y for all banks. A small number of observations were excluded if either assets or total costs were reported as zero, if input prices could not be constructed, or if constructed input prices were unrealistically large or small.

16. Because our operating cost measure excludes interest expense, we do not include either financial capital or the price of borrowed funds in the cost function.

17. The $Z_j$ values are functions that rescale the $\ln Y_j$ and the $\ln W_m$ so that they fall on the interval [$.1*2\pi$ , $.9*2\pi$]. See Berger, Leusner, and Mingo (1994) for a derivation of, and a justification for, this truncation. Because of computational limitations, we included only the $j=k=l$ terms in the triple summations.

18. See Bauer (1990) for a review of SFA methods. We cannot use the distribution-free approach because it only generates long-run estimates of cost efficiency. We cannot use the thick cost frontier approach because it estimates cost efficiency only for groups of banks. We do not use data envelopment analysis because it is often difficult to compare the efficiency of individual institutions using this approach. Berger, Hunter, and Timme (1993) discussed the relative strengths and weaknesses of these approaches.

19. Stevenson (1980) has shown that the truncated normal distribution is more flexible than the typical assumption of a half normal distribution, which restricts the probability density of the cost inefficiency term to decrease monotonically across banks. As estimated here, cost inefficiency includes both technical and allocative inefficiencies.

20. Note that the expected value of lnU conditional on $\eta$ is an imperfect estimate of actual inefficiency because the SFA process does not completely disentangle inefficiency from random error. For example, an efficient bank that experiences bad luck will have a high residual. The SFA estimation process will attribute part of residual to randomness and the other part of residual to inefficiency, even if the residual is strictly due to bad luck.

21. Average X-EFF fell from .9466 to .8780 (statistically different from each other at the 1% level) when we imposed both the translog and the half normal restrictions on the cost model using 1994 data (the lower of

these two averages remains slightly higher than estimates found elsewhere in the literature). An F-test of whether the coefficients on the trigonometric terms were jointly equal to zero rejected the null hypothesis of a translog specification at the 1% level, consistent with banking studies by McAllister and McManus (1993), Berger, Leusner and Mingo (1994), and Mitchell and Onvural (1996). Results were similar using annual data from other years in our data set.

We also found statistically significantly lower levels of X-EFF when we imposed only the translog **or** the half normal restrictions on the cost model. When the translog was substituted for the Fourier-flexible form (but the truncated normal distribution was retained for lnU), average X-EFF fell from .9466 to .8989 (significantly different at the 1% level) using 1994 data. When the half normal was substituted for the truncated normal (but the Fourier-flexible functional form was retained), average X-EFF fell from .9466 to .9191 (significantly different at the 1% level) for 1994 data. Again, an F-test rejected the translog null hypothesis at the 1% level. Relative to the half normal model, the distribution of estimated lnU in the truncated normal model was shifted to the left. This leaves less mass in the truncated normal tail of lnU, which may explain in part why our estimates of inefficiency are smaller than in most existing studies that use the half normal.

Note that measured inefficiency need not have decreased when we improved the fit of the cost function by substituting the Fourier-flexible for the translog, although a decrease is very likely when the fit is improved. The theoretical assumptions underlying the SFA method are that lnV contains all of the measurement error including specification error, that lnU contains only inefficiency, and the two terms are independent. Improving the specification causes the variance of lnV to decrease, but this does not affect the variance of the theoretical lnU term, which by definition does not contain any specification error. However, the **estimates** of lnU could either increase or decrease on average. That is, as the location of the estimated efficient cost frontier becomes more accurate and the two-sided errors become smaller, the measured one-sided deviations from the efficient frontier could become either smaller or larger on average. For example, if adding Fourier terms had pushed down the estimated cost frontier in a certain region of the data, then the estimates of lnU for

firms in that region could have increased as the frontier moved further away from them -- all else equal, such a change may have increased average inefficiency for the entire sample.

22. These calculations reflect only the direct effect of changes in lagged $Y_i$ on current $Y_j$ because these direct effects are the most appropriate measures for the hypotheses we are testing. We exclude indirect effects -- such as the effect of lagged $Y_i$ on current $Y_j$ through other lagged values of the Y's -- from our calculations in order to avoid mixing the intertemporal effects associated with our four different hypotheses.

23. We select all subsamples used in the regressions on the basis of lagged values rather than contemporaneous values of the endogenous variables in order to reduce sample selection bias. For this subsample, we required banks to be relatively cost efficient in every year, rather than just in individual years, to avoid including banks that only appeared to be efficient in individual years because they were skimping. This sampling procedure ensures that we observed only long-run cost efficient banks -- long-run cost-efficient skimpers as well as long-run cost-efficient non-skimpers.

24. Perhaps surprisingly, when the NPL equation is estimated for the entire sample of banks, the sum of the lagged CAP coefficients is positive and significant (see Table 2). There may be several reasons for this population-wide result. If banks anticipate high nonperforming loans in the future, they may act to raise capital in advance to provide a cushion against the associated loan losses. Alternatively, banks with high capital may be relatively less concerned about insolvency risk, and hence are more likely to skimp on loan monitoring, underwriting, and control. Finally, bank owners may elect to hold high capital to encourage normally risk-averse bank managers to take greater portfolio risks, yielding greater future nonperforming loans.

25. A survivor bias in our data may be one of the reasons for this result. Because we use end-of-year data we only observe the low-capital banks that survived the year, not the low-capital banks that failed during that year. A total of 1,290 commercial banks failed from 1985 through 1994, or an average of 129 per year. The expected negative relationship between problem loans and equity capital was probably very strong in these banks, because they essentially lost all of their capital.