FEDERAL RESERVE BANK OF PHILADELPHIA

Ten Independence Mall
Philadelphia, Pennsylvania 19106-1574
(215) 574-6428, www.phil.frb.org

Working Papers

Research Department

WORKING PAPER NO. 96-2/R

BANK CAPITALIZATION AND COST: EVIDENCE OF SCALE ECONOMIES IN RISK MANAGEMENT AND SIGNALING

Joseph P. Hughes
Rutgers University

Loretta J. Mester
Federal Reserve Bank of Philadelphia
and
The Wharton School, University of Pennsylvania

Revised February 1997

WORKING PAPER NO. 96-2/R

BANK CAPITALIZATION AND COST: EVIDENCE OF SCALE ECONOMIES IN RISK MANAGEMENT AND SIGNALING

Joseph P. Hughes
Rutgers University

and

Loretta J. Mester
Federal Reserve Bank of Philadelphia
and
The Wharton School, University of Pennsylvania

Revised February 1997

Forthcoming in *The Review of Economics and Statistics*.

The views expressed here are ours and do not necessarily represent the views of the Federal Reserve Bank of Philadelphia or of the Federal Reserve System.

We thank our referees for helpful comments and Avi Peled for research assistance.

Please address correspondence to Joseph P. Hughes, Department of Economics, Rutgers University, New Brunswick, NJ 08903, phone: (908) 932-7517, email: jphughes@rci.rutgers.edu or to Loretta J. Mester, Research Department, Federal Reserve Bank of Philadelphia, Ten Independence Mall, Philadelphia, PA 19106, phone: (215) 574-3807, email: Loretta.Mester@PHIL.frb.org.

**BANK CAPITALIZATION AND COST:
EVIDENCE OF SCALE ECONOMIES IN RISK MANAGEMENT AND SIGNALING**

Abstract

We amend the standard cost model to account for financial capital's role in banking. The cost function is conditioned on the level of capital, but we model the demand for financial capital so it can serve as a cushion against insolvency for potentially risk-averse managers and as a signal of risk for less informed outsiders. Scale economies are then computed without assuming that the bank chooses a level of capitalization that minimizes cost. We find evidence of substantial scale economies and that bank managers are risk averse and use the level of financial capital to signal the level of risk.

BANK CAPITALIZATION AND COST:
EVIDENCE OF SCALE ECONOMIES IN RISK MANAGEMENT AND SIGNALING

I. Introduction

Geographic restrictions on branching have contributed to the proliferation of banks in the U.S.—over 14,000 in 1982 and still over 9000 in 1997. The bank merger wave that began in the 1980s followed many states' liberalization of these restrictions and strongly suggests scale economies at work. An average of 190 mergers per year occurred from 1960 to 1982. This increased to an average of 423 mergers per year from 1980 to 1994, for a total of 6347 mergers in this latter period, representing 43 percent of all U.S. banks in 1980 (Rhoades, 1996). The largest U.S. bank mergers have also occurred since deregulation. In 1995 alone, 14.7 percent of large U.S. banks whose assets exceed $10 billion agreed to consolidate (Siems, 1996). The merger of Chase Manhattan and Chemical banks, announced in 1995, represents 8.7 percent of total U.S. bank assets (Rhoades, 1996).

Almost as often as bankers cite significant scale economies among their rationales for merging, economists complete econometric analyses of banking technology that fail to uncover evidence of these economies. (Mester, 1994 reviews this literature.) Studies that include banks of all sizes usually find slightly increasing returns to scale among small banks and slightly decreasing returns at large banks. Although there is still controversy over such fundamental issues as what constitutes a bank's outputs and inputs, a remarkable diversity of models has nevertheless reached the consensus that there are no important scale economies in U.S. banking.[1] This result is as robust as it is puzzling, given the recent merger wave.

Although the business of banking involves assessing, transforming, and managing a variety of risks in which banks have a comparative advantage, most studies of bank costs ignore risk, and this omission may account for the scale economy puzzle. In banking, greater size implies the potential for improved diversification, while improved diversification offers less risk and, hence, cost-savings in managing risk (Diamond, 1984) and in signaling the bank's riskiness to outsiders. Of course, a larger, more diversified bank may not accept the offer of less risk. Instead, it may exploit the reduced marginal cost of risk provided by larger scale and take on more risk. In taking more risk, it may "use up" these cost savings and conceal the potential economies that follow from scale-related diversification. Hence, if larger banks respond to a reduced marginal cost of risk by taking on and managing more risk, they may appear to have constant or even decreasing returns to scale because the extra risk is costly. Solving the scale economy puzzle may be as simple as accounting for risk.

1.1 Risk Signaling and Scale Economies. Banks, as delegated investors for their depositors, produce risky, illiquid assets (loans), while transforming them into safe, liquid claims (demand deposits) that their creditors can use to make payments. Demand deposits give banks specialized, private information they use to evaluate credit applications and to monitor loans. (Bhattacharya and Thakor, 1993 and Thakor, 1996 survey the literature.) Demandable debt also disciplines bank risk-taking. The information asymmetry that gives banks an edge in lending also gives uninsured depositors the incentive to monitor loan quality and to price its risk, and, in the extreme, to run on banks when insolvency seems imminent. Thus, banks' demand deposits generate liquidity risk. It is important that the loan portfolio appear to be of high quality to these depositors, since the actual credit risk is opaque to outsiders.

Credit risk and liquidity risk shape banking technology and define the roles of financial (equity) capital in preventing episodes of financial distress. Not only is financial capital a source of loanable funds, it is also a cushion for loan losses and, hence, protects the bank from the threat of insolvency that leads to regulatory intervention, liquidity crises, and, in the extreme, to loss of the bank's valuable charter.

Given a bank's scale and its inherent asset quality, an increase in financial capital reduces the probability of insolvency and provides an incentive to allocate additional resources to manage risk to protect the larger equity stake. Since financial capital constitutes the bank's own bet on its management of risk, it conveys a *credible* signal to depositors of the resources allocated to preserving capital and insuring the safety of their deposits.[2] Thus, higher levels of capitalization, given observable scale and inherent asset quality, inferred from measures such as the level of nonperforming loans, signal greater safety to depositors and, thus, reduce the probability of a liquidity crisis.

As a bank's scale increases, its loan portfolio and deposit base become more diversified. Hence, the same degree of protection against financial distress can be attained at a lower capital-to-asset ratio, while managing risk and protecting capital can be accomplished with relatively fewer resources. If we take into account capital's role as a signal of risk, diversification also reduces the cost of signaling, since it lowers the cost of protecting capital. In addition, it reduces the level of signal required, since scale and its implied degree of diversification are observable to outsiders.

If there are overall scale economies that result in part from these economies of risk management and risk signaling, then detecting them will be complicated by the reduced marginal cost of managing and signaling risk that is implied by better diversification. The reduced marginal cost of risk provides an incentive to take more risk, i.e., to reduce asset quality for a higher expected return. In turn, the lower asset quality may require more resources to manage the higher risk. The additional cost of these resources would reduce measured scale economies if no account were taken of the change in asset quality. Thus, in measuring overall scale economies, it is necessary to control not only for the output vector but also for output quality.

1.2 Modeling Strategy. Accounting for risk may be as simple as accounting for asset quality and capitalization, since the level of financial capital, given asset size and quality, is fundamentally linked to the bank's risk-management and risk-signaling. Accounting for capitalization and asset quality may, in turn, be as simple as conditioning the cost function on the level of capitalization and measures of asset quality. If the bank's assets are enumerated in the output vector, then including the level of capitalization captures the capital-to-asset ratio. Including measures of asset quality controls for a potentially endogenous source of risk that is influenced by the marginal cost of risk management. However, to compute scale economies, the level of capitalization must be made endogenous.

The standard procedure of computing scale economies from a conditional (or variable) cost function assumes that cost is minimized with respect to the conditional inputs. However, in this case, the conditioning level of capitalization serves as a cushion against loan losses and a signal of risk. Its level directly influences the probability of financial distress. Hence, there may be no reason to assume that the cost-minimizing level of capitalization is the appropriate risk cushion or risk signal.

Assuming a bank minimizes cost when it organizes production is equivalent to assuming it is risk neutral. But if the level of capital influences and signals the potential for costly financial distress, it may also signal nonneutrality toward risk. Risk-averse managers may find the level of risk implied by the least-cost level of capitalization to be unacceptable. If so, their choice of capital must be modeled by a broader objective than cost minimization.

Choosing a level of financial capital that reduces risk but increases cost amounts to trading profit for reduced risk. If managers are risk neutral, they will make no such trade. Hence, they maximize profit or, equivalently, minimize cost. Their utility is a function solely of profit. However, to allow for risk aversion, the manager's utility function must include other arguments in addition to profit.

In the sections that follow, we embed the problem of cost-minimization, conditioned on the level of financial capital, in a model of managerial utility-maximization that allows managers to increase cost to reduce risk. The utility-maximizing demand for capitalization obtained from this formulation does not necessarily minimize cost.

In fact, when managers are not risk-neutral, the reduced-form variable cost function that results from substituting the demand for financial capital into the variable cost function conditioned on capital will contain, not just the usual arguments of input prices and outputs, but also arguments of the revenue function. Thus, an explicit test for the bank managers' objective function is obtained: if variable cost depends on components of revenue, bank managers are acting in a nonneutral manner toward risk, taking into account both return and risk; if variable cost is independent of components of revenue, then managers' behavior is consistent with risk neutrality and cost minimization.

Scale economies can then be computed without assuming that the level of capitalization minimizes cost. Hence, a wider range of signals and cost configurations can be accommodated. Moreover, conditioning the cost function on measures of loan quality controls for an important source of endogenous risk that is influenced by scale-related diversification. Our specification should allow us to detect scale economies that would otherwise have been obscured by banks' assuming additional risk in response to improved diversification.

1.3 The Data and Evidence. Using 1989 and 1990 data on U.S. banks with assets over $1 billion, our estimates provide evidence that banks in all four size quartiles exhibit behavior inconsistent with risk-neutrality. In addition, we find evidence that financial capital serves as a signal of risk. Higher levels of capitalization are associated with higher variable costs of producing the given asset portfolio. In other words, the higher the capital-to-asset ratio, the higher the variable cost. This result may appear to be counterintuitive. Financial capital is a conditional argument in the variable cost function, and since it is a source of loanable funds, it is intuitively thought of as an input in the production of bank assets. Thus, it might be expected that variable cost, conditioned on the level of capital, would *be lower* with higher capitalization, since more capital allows a reduction in other funding sources.[3] However, the increase in variable cost that we find when financial capital increases means that greater amounts of the variable inputs are used when the level of capitalization is higher—in other words, *on net*, financial capital appears to behave more like an output than an input. This result is consistent with the signaling role of financial capital: when the level of financial capital is higher, more resources appear to be allocated to managing risk and preserving capital. We confirm this implication and also find that the marginal cost of signaling is significantly smaller for the largest quartile of banks.

In addition, we find that capitalization increases less than proportionally with assets and that there are significant scale economies at even the largest banks in the sample. Hence, by incorporating financial capital into the cost function in a manner that allows for its role in protecting against financial distress and in signaling risk, and that controls for sources of endogenous risk-taking, we have uncovered scale economies that have eluded the standard analysis.

2. Bank Production and Cost

The bank's technology is summarized by the transformation function $T(\mathbf{y},\mathbf{q},\mathbf{x},u,\mathbf{k}) = 0$, where $\mathbf{y}$ is a vector of quantities of outputs; $\mathbf{q}$ is a vector of variables characterizing output quality; u is uninsured deposits; $\mathbf{x}$ is a vector of inputs other than u; and $\mathbf{k}$ is a vector consisting of k_1, debt-based financial capital (subordinated debt), and k_2, equity-based financial capital and loan-loss reserves. $T(\mathbf{y},\mathbf{q},\mathbf{x},u,\mathbf{k})$ describes the bank's production possibilities set.

We assume banks are price-takers in the markets for inputs included in $\mathbf{x}$ so that the corresponding price vector $\mathbf{w}$ is competitively determined. The price of uninsured deposits, w_u, is also assumed to be competitively determined; however, if it includes a risk premium, then it may be affected by the bank's risk of failure as reflected in: (1) the quality, $\mathbf{q}$, of its outputs, (2) its capitalization, $\mathbf{k}$, relative to its size and mix of outputs, $\mathbf{y}$, and (3) a vector, $\boldsymbol{\theta}$, of variables that characterize riskiness but do not affect the transformation function. Thus, let $w_u = \omega F(\mathbf{y},\mathbf{q},\mathbf{k},\boldsymbol{\theta})$, where ω is a competitively determined, risk-free interest rate and $F(\mathbf{y},\mathbf{q},\mathbf{k},\boldsymbol{\theta}) \geq 1$ represents the risk premium. The cost of production is defined by,

$$C(\mathbf{y},\mathbf{q},\mathbf{w},\omega,\mathbf{k},\boldsymbol{\theta}) \equiv \min_{\mathbf{x},u} \left[\mathbf{w}\cdot\mathbf{x} + \omega F(\mathbf{y},\mathbf{q},\mathbf{k},\boldsymbol{\theta})u : T(\mathbf{x},u;\mathbf{y},\mathbf{q},\mathbf{k}) = 0\right]. \tag{1}$$

Conditioning cost on the bank's financial capital, $\mathbf{k}$, acknowledges that financial capital influences cost, since it provides an alternative to deposits as a source of loanable funds and it serves as a cushion against loan losses and a signal of risk. In addition, since the *level*, and not the price, is used in the cost function, it does not assume that the level of financial capital is cost-minimizing. This allows for the possibility that the bank is not risk neutral. Conditioning the cost function on measures of loan quality accounts for the endogenous element of risk that responds to scale-related diversification.

The formulation in equation (1) exhibits all the standard properties of a cost function. It should be noted, however, that this is a partially reduced-form cost function, obtained by substituting into the structural cost function the expression, $\omega F(\mathbf{y},\mathbf{q},\mathbf{k},\boldsymbol{\theta})$, for w_u, the price of uninsured deposits. Since the price of uninsured deposits depends on output levels and qualities, capitalization, a risk-free interest rate, and other risk variables, $\boldsymbol{\theta}$, this price does not explicitly appear in the cost function.[4]

3. The Demand for Financial Capital: Optimal Signaling

Financial capital provides a buffer against default and a signal of risk. Managers who are not risk neutral guard against financial distress and signal their bank's safety by choosing a level of capitalization that is likely to be greater than the cost-minimizing level. Hence, they trade profit for reduced risk. To allow for the possibility that managers are not risk neutral, we assume that management has well-behaved preferences defined, not simply in terms of the level of profit, but also in terms of the level of financial capital, relative to its given mix of assets, **y**, and their quality characteristics, **q**. Hence, if managers trade profit for reduced risk, their preference ordering of this trade-off depends on the size, composition, and quality of the bank's asset portfolio. The optimal level of capitalization may not be cost-minimizing; it may depend on management's degree of risk aversion and the size and specific characteristics of its portfolio of assets. Before completing the formulation of the managerial utility function, we consider the definition of profit.

To calculate profit, the cost of capital, $v_1k_1 + v_2k_2$, must be added to cost in equation (1). The bank's rate of interest on debt-based capital, v_1, consists of a risk-free component, γ_1, and a risk premium, $\Gamma_1(\mathbf{y},\mathbf{q},\mathbf{k},\boldsymbol{\theta}) \geq 1$; hence, $v_1 = \gamma_1\Gamma_1(\mathbf{y},\mathbf{q},\mathbf{k},\boldsymbol{\theta})$. Similarly, the required rate of return on equity-based capital is $v_2 = \gamma_2\Gamma_2(\mathbf{y},\mathbf{q},\mathbf{k},\boldsymbol{\theta})$. The interest rate charged on the i^{th} output is p_i. Revenue other than $\mathbf{p}\cdot\mathbf{y}$ is represented by $m \equiv p_mM$, where $p_m = 1$, so this "price" can be used to normalize profit when it is an argument of the utility function. Consequently, a proportional variation in all prices will not alter normalized profit and, hence, utility. In other words, normalized profit is homogeneous of degree zero in $(\mathbf{p},p_m,\mathbf{w},\omega,\gamma_1,\gamma_2)$ (or, equivalently, in $(\mathbf{p},m,\mathbf{w},\omega,\gamma_1,\gamma_2)$). Note that this means that the demand for financial capital will be homogeneous of degree zero in $(\mathbf{p},m,\mathbf{w},\omega,\gamma_1,\gamma_2)$ as well—a proportional variation in prices has no effect on normalized profit, which is an argument in the utility function; hence it will have no effect on the utility-maximizing level of financial capital.

Economic profit is defined by,

$$\hat{\pi} \equiv \mathbf{p}\cdot\mathbf{y} + m - C(\mathbf{y},\mathbf{q},\mathbf{w},\omega,\boldsymbol{\theta},\mathbf{k}) - v_1k_1 - v_2k_2, \tag{2}$$

while accounting profit, or net income, is defined by,

$$\pi \equiv \hat{\pi} + v_2k_2 = \mathbf{p}\cdot\mathbf{y} + m - C(\mathbf{y},\mathbf{q},\mathbf{w},\omega,\boldsymbol{\theta},\mathbf{k}) - v_1k_1. \tag{3}$$

The bank's return on equity is π/k_2, which consists of its required return, v_2, and its economic rent, $\hat{\pi}/k_2$.

The bank's preferences for profit, capitalization, quality, and size can be characterized by a twice-continuously differentiable, quasi-concave utility function, $U(\hat{\pi}/p_m,\mathbf{k},\mathbf{y},\mathbf{q})$. Utility is maximized with respect to profit and capitalization, subject to the budget constraint equation (2) and the vectors of outputs, **y**, and quality characteristics, **q**. If the bank maximizes profit, then **k**, **y**, and **q** will have no marginal effect on utility except through their effect on profit. On the other hand, if bank managers are risk averse, then they trade off profit for greater capitalization and hence less insolvency risk.

The bank's use of financial capital is subject to regulatory review; therefore, we must incorporate constraints defining the minimum capital adequacy level imposed by regulation. For generality, assume these constraints are functions of the bank's asset mix, size, and other parameters, $\boldsymbol{\tau}$. If the output vector is assumed to contain the bank's assets as some or all of its components, then we can denote the constraints by $k_i \geq G_i(\mathbf{y},\boldsymbol{\tau})$.

The demands for debt-based and equity-based financial capital and for economic profit follow from the solution to the problem,

$$\max_{\hat{\pi},\mathbf{k}} \; U(\hat{\pi}/p_m,\mathbf{k};\mathbf{y},\mathbf{q}) \tag{4}$$

$$\text{s.t.} \quad \sum_i p_i y_i + m - C(\mathbf{y},\mathbf{q},\mathbf{w},\omega,\boldsymbol{\theta},\mathbf{k}) - \gamma_1 g_1(\mathbf{y},\mathbf{q},\mathbf{k},\boldsymbol{\theta})k_1 - \gamma_2 g_2(\mathbf{y},\mathbf{q},\mathbf{k},\boldsymbol{\theta})k_2 - \hat{\pi} = 0$$

$$\text{and} \quad k_j \geq G_j(\mathbf{y},\boldsymbol{\tau}), \quad j = 1,2.$$

Assuming an interior solution, but not necessarily a binding regulatory constraint, and letting $\mathbf{z} \equiv (\mathbf{q},\boldsymbol{\theta},\boldsymbol{\tau},\mathbf{p},\mathbf{w},\omega,\boldsymbol{\gamma})$, where **q**, $\boldsymbol{\theta}$, $\boldsymbol{\tau}$, **p**, **w**, and $\boldsymbol{\gamma}$ are vectors, the demand functions for the components of financial capital are $\mathbf{k^m}(\mathbf{y},\mathbf{z},m) \equiv (k_1^m(\mathbf{y},\mathbf{z},m),k_2^m(\mathbf{y},\mathbf{z},m))$, and the demand function for economic profit is $\hat{\pi}^m(\mathbf{y},\mathbf{z},m)$. (The argument p_m is suppressed.) Substituting these demand functions into equation (3) yields the demand for net income, $\pi^m(\mathbf{y},\mathbf{z},m)$.

4. Cost with Endogenous Financial Capital: Tests for Risk Neutrality and for Signaling

Substituting the demand functions for financial capital, $\mathbf{k}^{\mathbf{m}}(\mathbf{y},\mathbf{z},m)$, into the cost function (1) yields cost that is unrestricted by capital,

$$C^{m}(\mathbf{y},\mathbf{q},\mathbf{z},m) \equiv C(\mathbf{y},\mathbf{q},\mathbf{w},\omega,\boldsymbol{\theta},\mathbf{k}^{\mathbf{m}}(\mathbf{y},\mathbf{z},m)). \tag{5}$$

When managers are risk-neutral and the capital constraint is not binding, the utility-maximizing demands for capital are identically equal to the cost-minimizing demands, and the cost function resulting from their substitution is an ordinary minimum cost function. On the other hand, when managers are not risk-neutral, their demands for capital are influenced by components of revenue as well as by input prices and the output levels, so that the cost function that follows from their substitution is, in turn, influenced by components of revenue.[5] Hence, testing for the presence of statistically significant revenue effects in the cost function provides evidence of the risk preferences of managers and of the potential signaling function of financial capital.

Stronger evidence of capital's signaling function is provided by the relationship of production and cost to the level of capitalization. As a source of loanable funds, capital is an input that can substitute for deposits and other types of borrowed funds. As a cushion against loan losses, capital protects the bank from financial distress and signals the bank's safety to less informed uninsured depositors and other outsiders. The higher the level of capitalization, given total assets and asset quality, the lower the risk of insolvency and the greater the degree of protection from financial distress. The signal given by a higher capital-to-asset ratio follows from the incentive to control moral hazard when more capital is at stake. If betting more financial capital gives banks an incentive to control risks better (or gives shareholders more incentive to monitor bank managers to ensure that they control risks better), then banks are likely to devote more labor and physical capital to production. The level of financial capital relative to total assets is apparent to outsiders, as is the level of financial capital relative to the amount of nonperforming loans. But, to assess the probability of insolvency, outsiders must also observe the degree of risk implied by the apparent loan quality, diversification, and resources allocated to managing risk. Since the amount of labor and physical capital devoted to risk management cannot be *directly* observed by outsiders, it must be inferred from the level of financial capital the bank chooses. To signal, then, the bank must incur the cost of increasing its level of capital.

If signaling is costly and capital is used for signaling, then the variable cost of producing any given vector of assets, conditioned on the level of financial capital and asset quality, should increase with the level of financial capital, i.e., the partial derivative of equation (1) with respect to financial capital should be *positive*. However, if financial capital is viewed as an input in its role as a source of loanable funds, then a higher level of capital permits a reduction in other sources of funds and, hence, a reduction in variable cost conditioned on the level of financial capital and asset quality, i.e., the partial derivative of equation (1) with respect to financial capital should be *negative*. Thus, when financial capital's role as a signal of risk dominates its role as a source of loanable funds, its marginal cost is positive.

5. The Empirical Implementation

Estimation of the cost function (5) relies on the (conditional) variable cost function (1), which is conditioned on the characteristics of quality and on the level of financial capital. In principle, this variable cost function is estimated jointly with its share equations and with a system of utility-maximizing demand equations for profit, debt-based financial capital, and equity capital. In the estimation described below, we use a simpler formulation to meet the limitations imposed by the data and the number of parameters.

We estimate the following model:

$$C(\mathbf{y},\mathbf{q},\mathbf{w},\omega,\mathbf{k},\boldsymbol{\theta}) \equiv \min \left[\mathbf{w}\cdot\mathbf{x} + \omega F(\mathbf{y},\mathbf{q},\mathbf{k},\boldsymbol{\theta})u:\ T(\mathbf{x},u;\mathbf{y},\mathbf{q},\mathbf{k}) = 0\right] \tag{6}$$

$$S_j \equiv \frac{w_j x_j}{C} = \frac{w_j}{C}\frac{\partial C}{\partial w_j}, \qquad w_j \in \mathbf{w} \tag{7}$$

$$\mathbf{k} \equiv \mathbf{k}^{\mathbf{m}}(\mathbf{y},\mathbf{z},m) \quad \text{where} \quad \mathbf{z} \equiv (\mathbf{q},\boldsymbol{\theta},\boldsymbol{\tau},\mathbf{p},\mathbf{w},\omega,\gamma), \tag{8}$$

making the following simplifications.[6] First, since nearly half of the banks in our sample do not have subordinated debt, the two types of financial capital are aggregated. Second, since estimating the behavior of costs does not require estimating the demand equation for profit, the profit function is dropped. Other simplifications and details are discussed below.

5.1 Data. We used 1989 and 1990 data from the quarterly Consolidated Reports of Condition and Income (the Call Reports). The sample of 286 banks includes all the U.S. banks that operated in branch-banking states and that reported over $1 billion in assets as of 1988Q4 and excludes the special-purpose Delaware banks chartered under that state's Financial Center Development Act and Consumer Credit Bank Act.[7] The banks included in the sample ranged from $1 billion to $74 billion in total assets in 1990.[8]

Our main specification includes five outputs: y_1 = real estate loans, y_2 = business loans (i.e., commercial and industrial loans, lease financing receivables, and agricultural loans), y_3 = loans to individuals, y_4 = other loans, and y_5 = securities, assets in trading accounts, fed funds sold and securities purchased under agreements to resell, and total investment securities. Each y_i is measured as the average of its dollar amount at the end of 1990 and its dollar amount at the end of 1989. Our alternative specification differs only from our main specification in that it includes a sixth output, y_6 = credit-risk equivalent amount of off-balance-sheet items (e.g., commitments, letters of credit, derivatives, etc.) at year-end 1990.[9,10] The price associated with this output is assumed to be the same across banks.

Since 138 banks in the sample have no subordinated debt, we aggregate the two categories of financial capital; hence, k is the average volume of equity capital, loan-loss reserves, and subordinated debt in 1990.

Four inputs, in addition to uninsured deposits and financial capital, are considered: (1) labor, (2) physical capital, (3) insured deposits, and (4) other borrowed money.[11] The corresponding input prices are: w_1 = salaries and benefits paid in 1990 ÷ average number of employees in 1990, w_2 = occupancy expense in 1990 ÷ average dollar value of net bank premises in 1990,[12] w_3 = (interest paid on small deposits [i.e., under $100,000] in 1990 − service charges on deposits paid to the bank in 1990) ÷ average volume of interest-bearing deposits less CDs over $100,000 in 1990, w_4 = total expense of fed funds purchased, securities sold under agreements to repurchase, obligations to the U.S. Treasury, and other borrowed money in 1990 ÷ average volume of these types of funds in 1990.

The price of uninsured deposits, $w_u = \omega F(\mathbf{y},\mathbf{q},k,\boldsymbol{\theta})$, consists of a risk-free component, ω, and a risk premium $F(\mathbf{y},\mathbf{q},k,\boldsymbol{\theta})$. A bank-specific risk-free rate, ω, is a weighted average of yields on 3-month, 1-year, 5-year, and 10-year Treasury securities in 1990, where the weights are calculated as the proportion of the bank's large time deposits with the corresponding maturity. The risk premium is a function of variables that are in the variable cost function, so once it is substituted into the variable cost function (as described in section 2) it is not explicitly estimated. Instead, ω serves as the parametric component of the price.[13]

The price of financial capital, $v = \gamma\Gamma(\mathbf{y},\mathbf{q},k,\boldsymbol{\theta})$, which appears in the demand equation for financial capital (and not in the variable cost function) consists of a risk-free component, γ, and a risk premium, $\Gamma(\mathbf{y},\mathbf{q},k,\boldsymbol{\theta})$. This price is incorporated into the capital equation in the same way that the price of uninsured deposits was incorporated into the cost function. Since the risk premium, $\Gamma(\mathbf{y},\mathbf{q},k,\boldsymbol{\theta})$, is a function of variables in the capital function, once it is substituted in, it is not explicitly estimated. Instead, the price of financial capital is parameterized by γ, its risk-free component. But since γ has no maturity, it is assumed to be the same across banks; thus, it does not appear in the cross-section estimation.[14]

The prices of the outputs, $\mathbf{p}$, are measured by dividing total interest income derived from each output by the average dollar amount of assets that are accruing interest in the output category in 1990. The quality of assets might be measured quite directly by the risk premium on each type of asset. Unfortunately, the data do not permit this calculation. Therefore, we include one quality variable, q, measured as the average total volume of nonperforming loans, i.e., loans past due 30 days or more plus nonaccruing loans. (Note that q is inversely related to quality.)[15] We include one risk variable, θ, measured by the standard deviation of the bank's yearly net income from 1986-1990. Since the regulatory parameters, $\boldsymbol{\tau}$, are the same for all banks, they are omitted in the cross-section model.

Variable cost, C, in equation (6), is measured by the sum of salaries and benefits, occupancy expense, and (interest paid on insured and uninsured deposits net of service charges, and the expenses of fed funds purchased, securities sold under agreements to repurchase, obligations to the U.S. Treasury, and other borrowed money)×(total loans and securities/total earning assets). Note that the costs of servicing capital (dividends to shareholders and interest payments to bondholders) are not included in C.

Finally, in the main specification, m, which is revenue other than **p·y**, is proxied by total noninterest income less service charges. This proxy is likely to include some income from sources related to output (e.g., revenues from loan sales), but is the best proxy available for this specification of outputs given the Call Report data. In the alternative specification of outputs, m is measured by other noninterest income; again, the best available proxy for this specification of outputs.

Table 1 summarizes the data for our sample.

5.2 Functional Form. We use the translog functional form for the cost function given in equation (6), and we assume the demand for financial capital is log-linear. Thus, the model, which consists of the (conditional) variable cost function, the cost share equations for the four inputs other than uninsured deposits, and the demand equation for financial capital, is:

$$\begin{aligned} \ln C = {} & a_0 + \sum_i a_i \ln y_i + \sum_j b_j \ln w_j + \tfrac{1}{2}\sum_i \sum_j s_{ij} \ln y_i \ln y_j + \tfrac{1}{2}\sum_i \sum_j g_{ij} \ln w_i \ln w_j \\ & + \sum_i \sum_j d_{ij} \ln y_i \ln w_j + f_k \ln k + f_q \ln q + f_\theta \ln\theta + \tfrac{1}{2} r_{kk} \ln k \ln k + r_{kq} \ln k \ln q \\ & + r_{k\theta} \ln k \ln\theta + \tfrac{1}{2} r_{qq} \ln q \ln q + r_{q\theta} \ln q \ln\theta + \tfrac{1}{2} r_{\theta\theta} \ln\theta \ln\theta \\ & + \sum_j h_{ki} \ln k \ln y_i + \sum_i h_{qi} \ln q \ln y_i + \sum_i h_{\theta i} \ln\theta \ln y_i \\ & + \sum_j t_{kj} \ln k \ln w_j + \sum_j t_{qj} \ln q \ln w_j + \sum_j t_{\theta j} \ln\theta \ln w_j + b_\omega \ln\omega + \tfrac{1}{2} g_{\omega\omega} \ln\omega \ln\omega \\ & + \sum_j g_{j\omega} \ln w_j \ln\omega + \sum_i d_{i\omega} \ln y_i \ln\omega + t_{k\omega} \ln k \ln\omega + t_{q\omega} \ln q \ln\omega + t_{\theta\omega} \ln\theta \ln\omega + \varepsilon \end{aligned} \qquad (9)$$

$$S_j = b_j + \sum_i g_{ij} \ln w_i + \sum_i d_{ij} \ln y_i + t_{qj} \ln k + t_{qj} \ln q + t_{\theta j} \ln\theta + g_{\omega j} \ln\omega + \xi_j , \quad j = 1,\ldots,4 \qquad (10)$$

$$\ln k = A_0 + \sum_i A_i \ln y_i + \sum_j B_j \ln w_j + B_\omega \ln\omega + R_q \ln q + R_\theta \ln\theta + R_m \ln m + \sum_i R_i \ln p_i + \nu \qquad (11)$$

where $s_{ij} = s_{ji}$ and $g_{ij} = g_{ji}$ by symmetry, and

$b_\omega = 1 - \sum_j b_j$, $g_{i\omega} = -\sum_j g_{ij}$, $\forall i$, $g_{\omega\omega} = -\sum_j g_{j\omega}$, $d_{i\omega} = -\sum_j d_{ij}$, $\forall i$,

$t_{k\omega} = -\sum_j t_{kj}$, $t_{q\omega} = -\sum_j t_{qj}$, and $t_{\theta\omega} = -\sum_j t_{\theta j}$ by linear homogeneity.

C = (conditional) variable cost

y_i = quantity of output i, i=1,...,5 for main specification and i=1,...,6 for alternative specification

w_j = price of input j (other than uninsured deposits), j=1,...4

ω = bank-specific risk-free rate of interest

k = financial capital

q = quality

θ = risk

m = revenue other than **p·y**,

p_i = price of output i, i=1,...,5

S_j = j^{th} cost share, i.e., expenditures on input j divided by variable cost, j=1,...,4

ε, ν, ξ_j are normally distributed error terms, j=1,...,4

and all variables (except the shares) are normalized by their means.

We allow the correlation of error terms on the cost function, share equations, and financial capital equation to be nonzero for any bank, but we assume the correlation is zero across banks. Since *ln* k is an endogenous variable that appears in the cost and share equations, we use iterative three-stage least squares to estimate the model. All the exogenous variables in the model as well as their squares and cross-products are used as instruments. (The squares and cross-products are used, since the square of the endogenous variable, i.e., $(ln\ k)^2$, appears in the cost equation (see Kelejian, 1971 and Greene, 1993, p. 609).) The estimates we obtain are asymptotically equivalent to maximum likelihood estimates.

Once the model is estimated, reduced-form coefficients can be calculated by substituting the demand equation for financial capital (11) into the cost and share equations. These coefficients can then be used to calculate certain characteristics like scale economies taking into account the endogeneity of financial capital.

6. The Empirical Results

The statistics reported in Tables 3-5 are evaluated at the mean levels of the exogenous variables for banks in four size categories that correspond to the sample quartiles determined by total assets in 1990.[16] These categories are: assets ≤ \$1.77 billion, \$1.77 billion < assets ≤ \$3.22 billion, \$3.22 billion < assets ≤

\$6.72 billion, and assets > \$6.72 billion. We also report these statistics evaluated at the mean levels of the exogenous variables for the entire size range of banks. Since these cost statistics are nonlinear functions of the parameters, standard errors are approximated by expanding each statistic as a Taylor series, dropping terms of order 2 and higher, and using the standard variance formula for linear functions of estimated parameters.

6.1 Economies in Capitalization. The parameters in the demand equation for financial capital (shown in Table 2) give the elasticity of demand with respect to the variables in the demand function, since the equation is log-linear. If banks become more diversified as their assets increase, then a proportional expansion in assets requires a less-than-proportional increase in financial capital to maintain the same degree of protection from insolvency. The elasticity of financial capital with respect to a scaled increase in asset size and performance (i.e., holding the ratio of nonperforming loans-to-assets constant) is given by

$$\begin{aligned} \text{ELASTK} &\equiv \sum_i \frac{\partial \ln \mathrm{k}}{\partial \ln \mathrm{y}_i} + \frac{\partial \ln \mathrm{k}}{\partial \ln \mathrm{q}} \\ &= \sum_i \mathrm{A}_i + \mathrm{R}_q. \end{aligned} \tag{12}$$

The estimate of ELASTK is 0.7153 (with standard error 0.03246) for the main specification and 0.8005 (with standard error 0.03080) for the alternative specification, indicating that there are significant economies in capitalization.

For both specifications, the elasticities with respect to all input prices are negative—significantly so at the 5 percent level for physical capital (B_2) and other borrowed money (B_4) in both specifications and at the 10 percent level for labor (B_1) in the main specification and for labor (B_1) and insured deposits (B_3) in the alternative specification. Hence, financial capital serves as a significant complement to labor and physical capital, a result that is necessary but not sufficient to indicate a signaling role for financial capital.

The other significant elasticities are those with respect to nonperforming loans (R_q), other revenue (R_m), and the average contractual rate of interest on business loans (R_2), and on securities (R_5). An increase in nonperforming loans, given total assets, implies a significant increase in the demand for financial capital.

The fact that components of the revenue function, p_2 and p_5, have a significant effect on the demand for financial capital, given the output vector and level of nonperformance, suggests banks may not be simple cost minimizers, but rather that they may be trading higher cost for reduced risk. We next examine whether these revenue variables have a significant effect on cost and whether there is evidence of signaling.

6.2 Bank Managers' Objectives. Table 3 presents the elasticities of cost with respect to the revenue variables, taking into account the endogeneity of k. Holding output, **y**, and nonperformance, q, constant, the revenue variables significantly affect cost for banks in all four size categories in both output specifications. Thus, managers appear to be trading higher costs for other objectives. In the context of financial intermediation, this trade-off is consistent with nonneutrality toward risk.

6.3 The Role of Financial Capital as a Signal of Risk. Table 4 shows that, in both output specifications, the marginal cost of financial capital, $\frac{\partial C}{\partial k}$, is positive for all four size quartiles, and that the employment of labor and physical capital increases with the level of financial capital and, hence, with the capital-to-asset ratio. Thus, a higher capital-to-asset ratio implies a higher cost of producing any given vector of outputs. Financial capital is not, on balance, substituting for variable inputs. To the contrary, more financial capital requires more nonfinancial inputs and *higher (conditional) variable cost* to protect it. Financial capital, then, behaves like an output or a signal. The marginal cost of financial capital is the cost of the signal.

There is no significant difference in the cost of this signal for banks across the three smaller size quartiles, but the cost of the signal (controlling for output level and asset quality) is significantly less for banks in the largest quartile compared to that of the smaller banks in the sample. This scale economy in signaling is also reflected in larger banks' use of labor and physical capital: the increase in the use of labor and physical capital in response to an increase in capital is less for banks in the fourth size quartile than for smaller banks. The cost of signaling might be less because larger banks are better able to diversify their portfolios, so that an increase in capital can be protected with fewer additional resources. Moreover, diversification might lower these banks' need to signal.

6.4 Scale Economies. We have emphasized the importance of controlling for the level of financial capital and asset quality in order to control for the effects of diversification when measuring scale economies. Thus, our measure of scale economies takes into account the endogeneity of financial capital by substituting (11) into (9), and then considers the effect on cost of a proportionate variation in the levels of all outputs in **y** and in the level of nonperforming loans, q, to hold *relative* nonperformance constant. Essentially, we are measuring scale economies using C^m given in equation (5), holding constant asset quality (i.e., the ratio of nonperforming loans total assets). Consider a composite output quantity and output quality bundle, $\zeta^0 \equiv (\mathbf{y^0},\mathbf{q^0})$, and suppose $\zeta = t\boldsymbol{\zeta^0}$. Totally differentiating with respect to a scaled variation in **y** and q yields,

$$\frac{dC}{\left(\frac{dt}{t}\right)} = \sum_i \frac{\partial C}{\partial y_i} y_i + \sum_i \frac{\partial C}{\partial k}\frac{\partial k}{\partial y_i} y_i + \frac{\partial C}{\partial q} q + \frac{\partial C}{\partial k}\frac{\partial k}{\partial q} q \ , \tag{13}$$

so that the degree of multiproduct scale economies is given by,

$$\begin{aligned} \text{SCALE} \equiv \frac{C}{\frac{dC}{\left(\frac{dt}{t}\right)}} &= \frac{C}{\sum_i \frac{\partial C}{\partial y_i} y_i + \sum_i \frac{\partial C}{\partial k}\frac{\partial k}{\partial y_i} y_i + \frac{\partial C}{\partial q} q + \frac{\partial C}{\partial k}\frac{\partial k}{\partial q} q} \\ &= \frac{1}{\sum_i \frac{\partial \ln C}{\partial \ln y_i} + \sum_i \frac{\partial \ln C}{\partial \ln k}\frac{\partial \ln k}{\partial \ln y_i} + \frac{\partial \ln C}{\partial \ln q} + \frac{\partial \ln C}{\partial \ln k}\frac{\partial \ln k}{\partial \ln q}} \end{aligned} \tag{14}$$

where SCALE > 1 implies multiproduct economies of scale, and SCALE < 1 implies diseconomies.

Table 5 indicates that, for both output specifications, banks in all four size categories are operating with relatively large scale economies. This is consistent with the claims of large institutions involved in mergers that scale and scope economies have motivated their combinations.[17] However, most studies of banking costs that ignored financial capital have found essentially constant returns to scale, often slightly increasing for small banks and slightly decreasing for large banks (see Mester, 1994 for a review of the literature). Our results here show that when allowance is made for the role of financial capital as a signal of risk and as a cushion against loan losses for potentially risk-averse bank managers, the scale economies become empirically detectable.

7. Conclusions

With seemingly minor changes to the standard technique of measuring banking production, we have uncovered empirical evidence that points to the crucial role played by financial capital in banking and financial intermediation. We use a standard cost function, conditioned on the level of financial capital and asset quality, and we model the demand for financial capital so that it can serve as a cushion against insolvency for potentially risk-averse managers and as a signal of bank risk to less informed outsiders. Scale economies can then be computed without assuming that the bank's chosen level of capitalization minimizes cost. We find evidence that bank managers are risk averse and use the level of financial capital to signal risk. For any given vector of outputs, they increase the level of financial capital to control risk and employ additional amounts of labor and physical capital to improve risk management and to preserve capital. We find evidence that as banks grow larger in asset size they are able to economize on the use of financial capital and the cost of signaling decreases for the largest banks, perhaps because larger banks are better able to diversify their portfolios. We find that banks of all sizes enjoy significant scale economies. Our results seem to reconcile the disparity between the finding of constant returns to scale in previous studies of bank costs that ignored financial capital and assumed risk-neutral bank managers, and the recent wave of large bank mergers, which bankers claim are driven in part by scale economies.

Table 1.—Means of the Variables*

	Banks with Assets under $1.77 Billion	Banks with Assets Between $1.77 and $3.22 Billion	Banks with Assets Between $3.22 and $6.72 Billion	Banks with Assets over $6.72 Billion	All Banks
y_1† real estate loans	0.3645	0.6508	1.282	3.680	1.490
y_2† business loans	0.2484	0.4686	0.9247	4.809	1.606
y_3† loans to individuals	0.1864	0.3288	0.6582	1.520	0.6722
y_4† other loans	0.03579	0.07722	0.1691	1.609	0.4703
y_5† securities	0.3230	0.5542	0.9757	3.469	1.327
y_6† credit-risk equivalent amount of off-balance-sheet items	0.06325	0.1553	0.4621	6.161	1.710
w_1††† price of labor	29.08	31.18	32.57	39.58	33.09
w_2†† price of physical capital	0.3907	0.3792	0.3810	0.4333	0.3960
w_3†† price of insured deposits	0.06010	0.06012	0.05935	0.0603	0.05996
w_4†† price of other borrowed money	0.08091	0.07901	0.08517	0.1034	0.08709
ω bank-specific risk-free rate	0.07673	0.07704	0.07702	0.07748	0.07707
k† financial capital	0.1069	0.1867	0.3558	1.606	0.5618
q† nonperforming loans	0.03800	0.08069	0.1533	0.7020	0.2426
θ† risk ≡ std. dev. of net income	0.005438	0.01229	0.02552	0.1294	0.04300

Table 1.—(*Continued*)

	Banks with Assets under \$1.77 Billion	Banks with Assets Between \$1.77 and \$3.22 Billion	Banks with Assets Between \$3.22 and \$6.72 Billion	Banks with Assets over \$6.72 Billion	All Banks
m in main specification† **total noninterest income − service charges**	0.01808	0.03352	0.06969	0.3777	0.1242
m in alternative specification† **other noninterest income**	0.009350	0.01401	0.03523	0.1983	0.06422
p_1† **price of y_1**	0.1078	0.1112	0.1096	0.1073	0.1090
p_2† **price of y_2**	0.1135	0.1078	0.1099	0.0974	0.1072
p_3† **price of y_3**	0.1246	0.1182	0.1254	0.1246	0.1232
p_4† **price of y_4**	0.08268	0.09520	0.08417	0.0797	0.08547
p_5† **price of y_5**	0.08650	0.08476	0.08594	0.0934	0.08764
C† **variable cost**	0.09286	0.1641	0.3231	1.202	0.4442

*The means of all variables except y_6 and m in the alternative specification are calculated for 71 banks in the first and fourth size quartiles, 72 banks in the second and third size quartiles, and 286 banks overall. The means of y_6 and m in the alternative specification are calculated for 71 banks in each size quartile and 284 banks overall.

†in billions of dollars ††in dollars per dollar †††in thousands of dollars per employee

Table 2—Parameter Estimates for Capital Demand Function

Parameter	Estimate (Approx. Std. Error)	Parameter	Estimate (Approx. Std. Error)	Parameter	Estimate (Approx. Std. Error)
Main Specification (5 outputs)					
A_0	0.1116* (0.02245)	$B_1 = \frac{d\,ln\,k}{d\,ln\,w_1}$	−0.08868** (0.05289)	$R_q = \frac{d\,ln\,k}{d\,ln\,q}$	0.2081* (0.03118)
$A_1 = \frac{d\,ln\,k}{d\,ln\,y_1}$	0.02972 (0.02392)	$B_2 = \frac{d\,ln\,k}{d\,ln\,w_2}$	−0.05753* (0.02172)	$R_\theta = \frac{d\,ln\,k}{d\,ln\,\theta}$	−0.01334 (0.02108)
$A_2 = \frac{d\,ln\,k}{d\,ln\,y_2}$	0.2052* (0.02618)	$B_3 = \frac{d\,ln\,k}{d\,ln\,w_3}$	−0.02479 (0.05720)	$R_m = \frac{d\,ln\,k}{d\,ln\,m}$	0.2029* (0.01765)
$A_3 = \frac{d\,ln\,k}{d\,ln\,y_3}$	0.01835 (0.01998)	$B_4 = \frac{d\,ln\,k}{d\,ln\,w_4}$	−0.1745* (0.03334)	$R_1 = \frac{d\,ln\,k}{d\,ln\,p_1}$	−0.05108 (0.04829)
$A_4 = \frac{d\,ln\,k}{d\,ln\,y_4}$	0.06793* (0.01382)	$B_\omega = \frac{d\,ln\,k}{d\,ln\,\omega}$	−0.7103 (0.5454)	$R_2 = \frac{d\,ln\,k}{d\,ln\,p_2}$	−0.09352* (0.04729)
$A_5 = \frac{d\,ln\,k}{d\,ln\,y_5}$	0.1859* (0.02368)			$R_3 = \frac{d\,ln\,k}{d\,ln\,p_3}$	−0.002272 (0.03848)
				$R_4 = \frac{d\,ln\,k}{d\,ln\,p_4}$	0.01183 (0.009019)
				$R_5 = \frac{d\,ln\,k}{d\,ln\,p_5}$	0.3405* (0.06287)
Alternative Specification (6 outputs)					
A_0	0.1776* (0.02505)	$B_1 = \frac{d\,ln\,k}{d\,ln\,w_1}$	−0.07911** (0.04513)	$R_q = \frac{d\,ln\,k}{d\,ln\,q}$	0.2179* (0.03006)
$A_1 = \frac{d\,ln\,k}{d\,ln\,y_1}$	0.007742 (0.02350)	$B_2 = \frac{d\,ln\,k}{d\,ln\,w_2}$	−0.05432* (0.01794)	$R_\theta = \frac{d\,ln\,k}{d\,ln\,\theta}$	−0.006306 (0.02050)
$A_2 = \frac{d\,ln\,k}{d\,ln\,y_2}$	0.1707* (0.02759)	$B_3 = \frac{d\,ln\,k}{d\,ln\,w_3}$	−0.08292** (0.04955)	$R_m = \frac{d\,ln\,k}{d\,ln\,m}$	0.08112* (0.01151)
$A_3 = \frac{d\,ln\,k}{d\,ln\,y_3}$	0.06807* (0.01878)	$B_4 = \frac{d\,ln\,k}{d\,ln\,w_4}$	−0.1750* (0.02805)	$R_1 = \frac{d\,ln\,k}{d\,ln\,p_1}$	−0.03164 (0.04465)
$A_4 = \frac{d\,ln\,k}{d\,ln\,y_4}$	0.07025* (0.01447)	$B_\omega = \frac{d\,ln\,k}{d\,ln\,\omega}$	0.09813 (0.4442)	$R_2 = \frac{d\,ln\,k}{d\,ln\,p_2}$	−0.1165* (0.04102)
$A_5 = \frac{d\,ln\,k}{d\,ln\,y_5}$	0.1855* (0.02354)			$R_3 = \frac{d\,ln\,k}{d\,ln\,p_3}$	0.01415 (0.03140)
$A_6 = \frac{d\,ln\,k}{d\,ln\,y_6}$	0.08035* (0.01750)			$R_4 = \frac{d\,ln\,k}{d\,ln\,p_4}$	0.006680 (0.007540)
				$R_5 = \frac{d\,ln\,k}{d\,ln\,p_5}$	0.4048* (0.05111)

Standard errors in parentheses.
*significant at 5% level **significant at 10% level

Table 3.—Elasticities of Cost With Respect to Revenue Variables†

	Banks with Assets under $1.77 Billion	Banks with Assets Between $1.77 and $3.22 Billion	Banks with Assets Between $3.22 and $6.72 Billion	Banks with Assets over $6.72 Billion	All Banks
Main Specification (5 outputs)					
$\frac{d\,ln\,C}{d\,ln\,p_1}$	−0.02175 (0.02093)	−0.02045 (0.01962)	−0.01975 (0.01886)	−0.01231 (0.01173)	−0.01437 (0.01372)
$\frac{d\,ln\,C}{d\,ln\,p_2}$	−0.03982** (0.02041)	−0.03743** (0.01911)	−0.03615** (0.01857)	−0.02253** (0.01208)	−0.02631** (0.01372)
$\frac{d\,ln\,C}{d\,ln\,p_3}$	−0.0009672 (0.01639)	−0.0009093 (0.01541)	−0.0008781 (0.01488)	−0.0005472 (0.009276)	−0.0006390 (0.01083)
$\frac{d\,ln\,C}{d\,ln\,p_4}$	0.005037 (0.003861)	0.004735 (0.003620)	0.004573 (0.003509)	0.002850 (0.002278)	0.003328 (0.002590)
$\frac{d\,ln\,C}{d\,ln\,p_5}$	0.1450* (0.03033)	0.1363* (0.02854)	0.1316* (0.02823)	0.08201* (0.02283)	0.09577* (0.02240)
Alternative Specification (6 outputs)					
$\frac{d\,ln\,C}{d\,ln\,p_1}$	−0.01824 (0.02588)	−0.01740 (0.02466)	−0.01644 (0.02321)	−0.01039 (0.01460)	−0.01225 (0.01730)
$\frac{d\,ln\,C}{d\,ln\,p_2}$	−0.06714* (0.02387)	−0.06405* (0.02264)	−0.06050* (0.02152)	−0.03824* (0.01443)	−0.04508* (0.01655)
$\frac{d\,ln\,C}{d\,ln\,p_3}$	0.008157 (0.01807)	0.007783 (0.01724)	0.007351 (0.01625)	0.004647 (0.01025)	0.005478 (0.01209)
$\frac{d\,ln\,C}{d\,ln\,p_4}$	0.003851 (0.004333)	0.003674 (0.004128)	0.003470 (0.003906)	0.002193 (0.002507)	0.002586 (0.002915)
$\frac{d\,ln\,C}{d\,ln\,p_5}$	0.2334* (0.03383)	0.2226* (0.03176)	0.2103* (0.03111)	0.1329* (0.02797)	0.1567* (0.02769)

†Cost statistics evaluated at size-category means of input prices, quality measure, risk measure, and output levels. Approximate standard errors in parentheses.

*significant at 1% level **significant at 10% level

p_1 = price of real estate loans p_2 = price of business loans p_3 = price of loans to individuals
p_4 = price of other loans p_5 = price of securities

Table 4.—Marginal Cost of Capital and Derivatives of Nonfinancial Inputs With Respect to Financial Capital†

	Banks with Assets under $1.77 Billion	Banks with Assets Between $1.77 and $3.22 Billion	Banks with Assets Between $3.22 and $6.72 Billion	Banks with Assets over $6.72 Billion	All Banks
Main specification (5 outputs)					
$\frac{\partial C}{\partial k}$	0.3249* (0.04064)	0.3074* (0.03640)	0.3162* (0.03901)	0.1854* (0.04269)	0.2004* (0.03301)
$\frac{\partial \text{labor}}{\partial k}$	0.005759* (0.0004790)	0.005321* (0.0004250)	0.005445* (0.0004412)	0.003673* (0.0004132)	0.004158* (0.0003881)
$\frac{\partial \text{ physical capital}}{\partial k}$	0.1820* (0.01446)	0.1862* (0.01425)	0.1985* (0.01534)	0.1479* (0.01533)	0.1526* (0.01345)
Alternative specification (6 outputs)					
$\frac{\partial C}{\partial k}$	0.4273* (0.04025)	0.4127* (0.03491)	0.4095* (0.03833)	0.2400* (0.04708)	0.2580* (0.03844)
$\frac{\partial \text{labor}}{\partial k}$	0.005233* (0.0004820)	0.004886* (0.0004264)	0.004908* (0.0004518)	0.003020* (0.0004260)	0.003416* (0.0004113)
$\frac{\partial \text{ physical capital}}{\partial k}$	0.1874* (0.01507)	0.1938* (0.01476)	0.2050* (0.01632)	0.1437* (0.01636)	0.1477* (0.01468)

†Cost statistics evaluated at size category means of input prices, quality measure, risk measure, and output levels. Approximate standard errors in parentheses.

*significantly different from 0 at 1% level

Table 5.—Scale Economies[†]

	Banks with Assets under \$1.77 Billion	Banks with Assets Between \$1.77 and \$3.22 Billion	Banks with Assets Between \$3.22 and \$6.72 Billion	Banks with Assets over \$6.72 Billion	All Banks
Main specification (5 outputs)					
SCALE	1.138[#] (0.02539)	1.147[#] (0.02135)	1.131[#] (0.02060)	1.148[#] (0.03019)	1.157[#] (0.02294)
Alternative Specification (6 outputs)					
SCALE	1.059[#] (0.02067)	1.077[#] (0.01708)	1.085[#] (0.01727)	1.145[#] (0.02809)	1.116[#] (0.02236)

[†]Cost statistics evaluated at size category means of input prices, quality measure, risk measure, and output levels. Approximate standard errors in parentheses.

[#]significantly different from 1 at 1% level

REFERENCES

Berger, Allen, and Robert DeYoung, “Problem Loans and Cost Efficiency in Commercial Banks,” *Journal of Banking and Finance* (1997).

Bhattacharya, Sudipto, and Anjan V. Thakor, “Contemporary Banking Theory,” *Journal of Financial Intermediation* 3 (1993), 2-50.

Diamond, D., “Financial Intermediaries and Delegated Monitoring, *Review of Economic Studies* 51 (1984), 393-414.

Diewert, W. E., "Duality Approaches to Microeconomic Theory," in K. J. Arrow and M. D. Intriligator (eds.), *Handbook of Mathematical Economics*, vol 2 (New York: North-Holland, 1982), 535-599.

Greene, William H., *Econometric Analysis*, second edition (New York: Macmillan Publishing Company, 1993).

Hancock, Diana, "The Financial Firm: Production with Monetary and Nonmonetary Goods," *Journal of Political Economy* 93 (1985), 859-880.

Hancock, Diana, "A Model of the Financial Firm with Imperfect Asset and Deposit Liabilities," *Journal of Banking and Finance* 10 (1986), 37-54.

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

Hunter, William C., Stephen G. Timme, and Won Keun Yang, "An Examination of Cost Subadditivity and Multiproduct Production in Large U.S. Banks." *Journal of Money, Credit, and Banking* 22 (1990), 504-525.

Jagtiani, Julapa, and Anya Khanthavit, “Scale and Scope Economies at Large Banks: Including Off-Balance Sheet Products and Regulatory Effects (1984-1991),” *Journal of Banking and Finance* 20 (1996), 1271-1287.

Kelejian, H, "Two-Stage Least Squares and Econometric Systems Linear in Parameters but Nonlinear in the Endogenous Variables," *Journal of the American Statistical Association* 66 (1971), 373-374.

Leland, Hayne, and David Pyle, "Informational Asymmetries, Financial Structure, and Financial Intermediation," *Journal of Finance* 32 (1977), 371-388.

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

Mester, Loretta J., "Testing for Expense Preference Behavior: Mutual Versus Stock Savings and Loans," *The RAND Journal of Economics* 20 (1989), 483-498.

Mester, Loretta J., "Agency Costs among Savings and Loans," *Journal of Financial Intermediation* 1 (1991), 257-278.

Mester, Loretta J., "How Efficient Are Third District Banks?" *Business Review*, Federal Reserve Bank of Philadelphia (January/February 1994), 3-18.

Rhoades, Stephen A., "Bank Mergers and Industrywide Structure, 1980-94," Staff Study 169, Board of Governors of the Federal Reserve System (January 1996).

Ross, S. A., "The Determination of Financial Structure: The Incentive Signalling Approach," *Bell Journal of Economics* (1977), 23-40.

Shaffer, S., “A Revenue-Restricted Cost Study of 100 Large Banks,” *Applied Financial Economics* 4 (1994), 193-205.

Siems, Thomas F., "Bank Mergers and Shareholder Wealth: Evidence from the 1995's Megamerger Deals," *Financial Industry Studies*, Federal Reserve Bank of Dallas (August 1996), 1-12.

Thakor, Anjan V., "The Design of Financial Systems: An Overview," *Journal of Banking and Finance* 20 (1996), 917-948.

NOTES

1. It should be noted that a few studies that estimated cost functions for large banks did find significant scale economies (see, e.g., Shaffer, 1994).

2. For the role of capital structure in signaling, see Ross (1977) and Leland and Pyle (1977).

3. An increase in the level of a conditional or fixed input, implies that the cost of the variable inputs used to produce a given output level would be lower, all else equal.

4. The structural cost function is a function of ($\mathbf{y}$,$\mathbf{q}$,$\mathbf{w}$,w_u,$\mathbf{k}$). Substituting w_u, which is a function of ($\mathbf{y}$,$\mathbf{q}$,$\mathbf{k}$,$\boldsymbol{\theta}$,ω), into the structural cost function yields the partially reduced-form conditional cost function, which is a function of ($\mathbf{y}$,$\mathbf{q}$,$\mathbf{w}$,ω,$\mathbf{k}$,$\boldsymbol{\theta}$). Note that since the structural cost function is linearly homogeneous in input prices, ($\mathbf{w}$,w_u), the partially reduced-form cost function is linearly homogeneous in ($\mathbf{w}$,ω). The fully reduced-form cost function is obtained by substituting the demand for capital, $\mathbf{k}$, into the partially reduced-form cost function. In Section 3, we show $\mathbf{k}$ is a function of arguments of the revenue function as well as ($\mathbf{w}$,ω). Hence, the fully reduced-form cost function will not, in general, be linearly homogeneous in ($\mathbf{w}$,ω).

5. For example, an increase in the ith output price, p_i, increases the revenue, $\mathbf{p}\cdot\mathbf{y}$ + m, that results from any given output vector, $\mathbf{y}$. The level of capital that minimizes cost, or equivalently, that maximizes the profit deriving from the given $\mathbf{y}$, is unaffected by the increase in p_i, so that the profit function, conditioned on financial capital and $\mathbf{y}$, is shifted up vertically in (k,π) space. The utility-maximizing capital choice of risk-neutral managers, whose indifference curves are flat in this space, would not change. But the capital choice of non-risk-neutral managers, whose indifference curves are convex, would change.

6. In the estimation we omit one of the share equations (that for uninsured deposits), since the shares must sum to one. The estimation is not affected by the choice of which share equation to delete, since maximum likelihood estimates are obtained.

7. We exclude banks in unit-banking states and the Delaware legislated banks to help control for

the regulatory environment. In addition, a few banks were dropped because of missing or misreported data.

8. It could be that banks with less than the regulatory required level of capital are unable to choose their capital levels, because it is being dictated by the regulators. The requirement in 1990 was capital to total assets of at least 6 percent. Only two banks in the sample had capital ratios less than 6 percent, and dropping these banks from the sample does not qualitatively change any of the results reported below. Under the risk-based capital standards, effective at the end of 1992, total capital to risk-adjusted assets was required to be at least 8 percent. Dropping banks with total capital to total assets under 8 percent (a stricter requirement) does not change our results.

9. We use the year-end l990 levels (as reported on Schedule RC-R of the Call Report) rather than the average of year-end 1989 and year-end 1990 levels because the Call Report specification of these items changed between 1989 and 1990. Because two banks reported credit-risk equivalent amounts equal to zero, the sample for the alternative specification includes 284 banks.

10. Three different ways have been used to measure off-balance-sheet items in the cost function literature: the credit-risk equivalent amounts (where the Basle Accord risk weights are used to convert the unweighted amounts of off-balance-sheet items to be risk-equivalent to loans), the unweighted amounts, and income from these activities less capital gains and losses on asset sales (see, e.g., Jagtiani and Khanthavit, 1996, which uses the first two measures). A drawback of the third measure is that it can be negative, which is undesirable for an output measure. Our results are robust to measuring off-balance-sheet items by their unweighted amounts.

11. Hughes and Mester (1993) developed a test to determine whether deposits are outputs or inputs. A variable cost (VC) function in which insured and uninsured deposits are entered in levels is estimated. If the derivative of VC with respect to insured deposits is positive, then insured deposits are an output; if the derivative is negative, then insured deposits are an input. The same test can be applied for uninsured deposits. Using a data set similar to the one employed here,

Hughes and Mester (1993) found that both types of deposits were inputs. Hence, we treat them as inputs here as well.

12. This measure of the unit price of physical capital has been used in many other cost studies, including Hughes and Mester (1993), Mester (1991), and Hunter, Timme, and Yang (1990). As an alternative, the rental cost per square foot of office space at the bank's headquarters location could be used. However, it is not clear this would be a better proxy, since many of the banks in the sample have many branches at various locations. While in theory one could use the average rental cost over all markets in which the bank operates, data on branch location were not available.

13. This procedure for dealing with an endogenous input price was suggested by Diewert (1982), pp. 584-585.

14. The coefficient on γ could be derived, however, since the demand for financial capital equation is homogeneous of degree zero in $(\mathbf{p},m,\mathbf{w},\omega,\gamma)$. Said differently, since γ does not explicitly appear, the homogeneity of degree zero condition places no restrictions on the coefficients on prices that do appear in the demand for capital equation.

15. This is an ex post measure of quality rather than an ex ante measure—not all low quality loans end up being nonperforming loans, and not all loans that are performing well today will continue to do so—but it seems to be the best available measure of the resources that went into monitoring the bank's loans. Also note that while the quantity of a bank's nonperforming loans will be influenced by the macroeconomy, its cross-sectional variation measures differences in quality across the banks, so long as macroeconomic factors are the same across regions. Berger and DeYoung (1997) discuss the inclusion of nonperforming loans in efficiency cost studies of banks.

16. The parameter estimates and goodness-of-fit measures for the main and alternative specifications are available from the authors.

17. Although not shown here to save space, we find that some evidence of product-specific scope economies.