

Hughes, Joseph P.; Mester, Loretta J.

Working Paper

Efficiency in banking: Theory, practice, and evidence

Working Paper, No. 2008-01

Provided in Cooperation with:

Department of Economics, Rutgers University

Suggested Citation: Hughes, Joseph P.; Mester, Loretta J. (2008) : Efficiency in banking: Theory, practice, and evidence, Working Paper, No. 2008-01, Rutgers University, Department of Economics, New Brunswick, NJ

This Version is available at:

<https://hdl.handle.net/10419/31302>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Efficiency in Banking: Theory, Practice, and Evidence

Joseph P. Hughes
Rutgers University

and

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

January 2008

Prepared for the *Oxford Handbook of Banking*

Abstract. Great strides have been made in the theory of bank technology in terms of explaining banks' comparative advantage in producing informationally intensive assets and financial services and in diversifying or offsetting a variety of risks. Great strides have also been made in explaining sub-par managerial performance in terms of agency theory and in applying these theories to analyze the particular environment of banking. In recent years, the empirical modeling of bank technology and the measurement of bank performance have begun to incorporate these theoretical developments and yield interesting insights that reflect the unique nature and role of banking in modern economies. This chapter gives an overview of two general empirical approaches to measuring bank performance and discusses some of the applications of these approaches found in the literature.

Keywords: Bank, efficiency, risk, cost, profit, agency costs, X-inefficiency

Correspondence to:

Mester at Research Department, Federal Reserve Bank of Philadelphia, Ten Independence Mall,
Philadelphia, PA 19106-1574; phone: 215-574-3807; fax: 215-574-4303; email:
Loretta.Mester@phil.frb.org.

Hughes at Department of Economics, Rutgers University, New Brunswick, NJ 08901-1248;
phone: 732- 932-7517; fax: 732-932-7416; email: jphughes@rci.rutgers.edu.

The views expressed here are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Philadelphia or of the Federal Reserve System.

Introduction

What do commercial banks do? What are the key components of banking technology? What determines whether banks operate efficiently? The literature on financial intermediation suggests that commercial banks, by screening and monitoring borrowers, can solve potential moral hazard and adverse selection problems caused by the imperfect information between borrowers and lenders. From the information obtained from checking account transactions and other sources, banks assess and manage risk, write contracts, monitor contractual performance, and, when required, resolve nonperformance problems. (Bhattacharya and Thakor (1993) review the modern theory of financial intermediation.)

Banks' ability to ameliorate informational asymmetries between borrowers and lenders and their ability to manage risks are the essence of bank production. These abilities are integral components of bank output and influence the managerial incentives to produce financial services prudently and efficiently. That banks' liabilities are demandable debt gives banks an incentive advantage over other intermediaries. The relatively high level of debt in a bank's capital structure disciplines managers' risk-taking and their diligence in producing financial services by exposing the bank to an increased risk of insolvency. The demandable feature of the debt, to the extent it is not fully insured, further heightens performance pressure and safety concerns by increasing liquidity risk. These incentives tend to make banks good monitors of their borrowers. Hence, the banking relationship can improve the financial performance of bank customers and increase access to credit for firms too informationally opaque to borrow in public debt and equity markets. The uniqueness of bank production, in contrast to the production of other types of lenders, is derived from the special characteristics of banks' capital structure: the funding of

informationally opaque assets with demand deposits.¹ (For a discussion of the optimal capital structure of commercial banks, see Calomiris and Kahn (1991) and Flannery (1994).)

But banks' ability to perform efficiently – to obtain accurate information concerning its customers' financial prospects and to write effective contracts and to enforce them – depends in part on the property rights, legal, regulatory, and contracting environments in which they operate. Such an environment includes accounting practices, chartering rules, government regulations, and the market conditions (e.g., market power) under which banks operate.

Differences in these features across political jurisdictions can lead to differences in the efficiency of banks across jurisdictions.² The operating environment can also influence the external and internal mechanisms that discipline bank managers. Internal discipline might be induced or reduced by organizational form, ownership and capital structure, governing boards, and managerial compensation. External discipline might be induced or reduced by government regulation and the safety net, capital market discipline (takeovers, cost of funds, stakeholders' ability to sell stock (stock price)), managerial labor market competition, outside blockholders (equity and debt), and product market competition.³

¹ Berlin and Mester (1999) find empirical evidence of an explicit link between banks' liability structure and their distinctive lending behavior. As discussed in Mester (2007), relationship lending is associated with lower loan rates, less stringent collateral requirements, a lower likelihood of credit rationing, contractual flexibility, and reduced costs of financial distress for borrowing firms. Banks' access to core deposits, which are rate inelastic, enable banks to insulate borrowers with whom they have durable relationships from exogenous credit shocks. Mester, Nakamura, and Renault (2007) also find empirical evidence of a synergy between the liability and asset sides of a commercial bank's balance sheet, showing that information on the cash flows into and out of a borrower's transactions account can help an intermediary monitor the changing value of collateral that a small-business commercial borrower has posted.

² Demirgüç-Kunt, Kane, and Laeven (2007) use a sample of 180 countries to study the external and internal political features that influence the adoption and design of deposit insurance, which, in turn, affect the efficiency of the domestic banking system.

³ LaPorta, Lopez-de-Silanes, and Shleifer (2002) examine banking systems in 92 countries and find that government ownership is correlated with poorer countries and countries with less developed financial systems, poorer protection of investors' rights, more government intervention, and poorer performance of institutions. They also find that government ownership is associated with higher cost ratios and wider interest rate margins. Aghion, Alesina, and Trebbi (2007) provide evidence that democracy has a positive impact on productivity growth in more advanced sectors of the economy, possibly by fostering entry and competition.

I. Banking Technology and Performance

I.A. The empirical measurement of banking technology and performance

There are two broad approaches to measuring technology and explaining performance: nonstructural and structural. Using a variety of financial ratios that capture various aspects of performance, the nonstructural approach compares performance among banks and considers the relationship of performance to investment strategies and other factors such as characteristics of governance. For example, the nonstructural approach might investigate technology by asking how performance ratios are correlated with such investment strategies as growing by asset acquisitions and diversifying or focusing the bank's product mix. It looks for evidence of agency problems in correlations of performance ratios and variables characterizing the quality of banks' governance. While informal and formal theories may motivate some of these investigations, no general theory of performance provides a unifying framework for these studies.

The structural approach is choice-theoretic and, as such, relies on a theoretical model of the banking firm and a concept of optimization. The older literature applies the traditional microeconomic theory of production to banking firms. The newer literature views the bank as a financial intermediary that produces informationally intensive financial services and diversifies risks, and combines the theory of financial intermediation with the microeconomics of bank production. This helps guide the choice of outputs and inputs in the bank's production structure. For example, as discussed in Mester (forthcoming), the standard application of efficiency analysis to banking does not allow bank production decisions to affect bank risk. This rules out the possibility that scale- and scope-related improvements in diversification could lower the cost of borrowed funds and induce banks to alter their risk exposure. Also, much of the earlier literature does not account for the bank's role in producing information about its borrowers in its

underwriting decisions when specifying the bank's outputs and inputs. An exception is Mester (1992), which directly accounted for banks' monitoring and screening role by measuring bank output treating loans purchased and originated loans as separate outputs entailing different types of screening, and treating loans held on balance sheet and loans sold as separate outputs entailing different types of monitoring.

Banks make choices about their capital structure and the amount of risk to assume, which should be taken into account when modeling bank production. Part of the input and output prices a bank faces are not exogenous – the bank makes strategic decisions regarding asset quality and capital structure, which affect the risk premium in its output and input prices. These decisions also relate to how one should view bank performance. In the standard efficiency literature, the bank is assumed to choose a production plan that minimizes costs given its output mix and input prices or that maximizes profits given the prices of its inputs and outputs. In newer research (e.g., Hughes, Lang, Mester, and Moon, 2000; Hughes, 1999; Hughes, Lang, Mester, and Moon, 1999; and Hughes, Mester, and Moon, 2001) bank managers are modeled as maximizing their utility, which is a function of market value and risk. To the extent that production decisions affect bank risk, they also affect the discount rate applied to evaluating the present value of costs and profit streams. Production decisions that increase expected profit but also increase the discount rate applied to that profit may not increase the bank's market value. In addition, managers may trade off expected return and risk, so that production choices that maximize managers' utility depend not only on the expected profits they generate but also on the variability of the profit stream they generate. Banks with high levels of agency problems between owners and managers might choose utility-maximizing production plans, but these need not be value-maximizing plans if the risk-return tradeoffs being made are not efficient.

How one gauges performance in structural models, then, depends on whether one views the bank as minimizing cost, maximizing profits, or maximizing managerial utility. In the latter case, one would want to gauge the trade-offs between risk and expected return being made in banks with minimum agency problems between owners and managers, i.e., banks with strong corporate controls (see Hughes, Mester, and Moon, 2001). In both the structural and nonstructural approaches, the performance metric and the specification of the performance equation reflect implicitly or explicitly an underlying theory of managerial behavior.

As a general specification of the structural and nonstructural approaches, let y_i represent the measure of the i^{th} bank's performance. Let z_i be a vector of variables that capture key components of the i^{th} bank's technology (e.g., output levels and input prices) and τ_i be a vector of variables affecting the technology (e.g., the ratio of nonperforming to total loans). Jensen and Meckling (1979) add a vector, θ_i , of characteristics of the property-rights system, contracting, and regulatory environment in which the i^{th} firm operates (e.g., whether the country has a deposit insurance scheme and the degree of investor protection) and a vector, ϕ_i , of characteristics of the organizational form and the governance and control environment of the i^{th} firm (e.g., whether the bank is organized as a mutual or stock-owned firm, the degree of product market concentration, and the number of outside directors on its board). When the sample of banks used in the estimation includes financial institutions located in environments with different property rights and contracting environments or with different governance and control structures, estimating this model permits one to investigate how these differences are correlated with differences in bank performance.

Allowing for random error, the performance equation to be estimated takes the form,

$$(1) \quad y_i = f(z_i, \tau_i, \phi_i, \theta_i | \beta) + \varepsilon_i.$$

The specification of the vectors z_i and τ_i differs between the structural and nonstructural approaches.

I.B. The structural approach to bank efficiency measurement: cost minimization, profit maximization, and managerial utility maximization

The *structural approach* usually relies on the economics of cost minimization or profit maximization, where the performance equation denotes a cost function or a profit function. Occasionally, the structural performance equation denotes a production function. While estimating a production function might tell us if the firm is *technically efficient*, i.e., if managers organize production so that the firm maximizes the amount of output produced with a given amount of inputs (so that the firm is operating on its production frontier), we are more interested in *economic efficiency*, i.e., whether the firm is correctly responding to relative prices in choosing its inputs and outputs, which subsumes technical efficiency.

In the newer literature, the optimization problem is managerial utility maximization, where the manager trades off risk and expected return. The vector z includes input prices and output prices in a profit function. In the cost function and the nonstandard profit function (Humphrey and Pulley, 1997), the vector contains input prices and output levels. In all of these cases, τ might include controls like nonperforming loans to total loans or off-balance-sheet assets to total assets.

These functions can also differ by the definition of cost they use: accounting (cash-flow) cost excludes the cost of equity capital, while economic cost includes it. The theoretically proper specification of accounting cost is addressed in section I.E. The challenge of specifying economic cost is estimating the cost of equity capital. McAllister and McManus (1993) arbitrarily pick the required return and assume it is uniform across banks. Clark (1996) and

Fiordelisi (2007) use the Capital Asset Pricing Model to estimate it. Fiordelisi (2007) describes the resulting profit function as “Economic Value Added.”

The structural performance equation can be fitted to the data as an average relationship, which assumes that all banks are equally efficient at minimizing cost or maximizing profit, subject to random error, ε_i , which is assumed to be normally distributed. Alternatively, the structural performance equation can be estimated as a *stochastic frontier* to capture best-practice and to gauge inefficiency, the difference between the best-practice performance and achieved performance. Berger and Mester (1997) review the estimation methods. Note that best-practice performance is sometimes called potential performance. However, this is somewhat of an abuse of terms since the best-practice performance does not necessarily represent the best possible practice, but merely the best practice observed among banks in the sample (see Berger and Mester, 1997, and Mester, forthcoming).

In the stochastic frontier, the error term, ε_i , consists of two components; one is a two-sided random error that represents noise (v_i), and one is a one-sided error representing inefficiency (μ_i). The stochastic frontier approach disentangles the inefficiency and random error components by making explicit assumptions about their distributions. The inefficiency component measures each bank’s extra cost or shortfall of profit relative to the frontier – the best practice performance observed in the sample.⁴ Let y_i denote either the cost or profit of firm i . The stochastic frontier gives the highest or lowest potential value of y_i given z_i , τ_i , ϕ_i , and θ_i ,

$$(2) \quad y_i = F(z_i, \tau_i, \phi_i, \theta_i | \beta) + \varepsilon_i,$$

⁴ Leibenstein (1966) called such inefficiency, which can result from poor managerial incentives or the failure of the labor market to allocate managers efficiently and to weed out incompetent managers, *X-inefficiency*. Jensen and Meckling (1976) called such inefficiency *agency costs* and provided a theoretical model of managerial utility maximization to explain how, when incentives between managers and outside stakeholders are misaligned, managers may trade off the market value of their firm to enjoy more of their own private benefits, such as

where $\varepsilon_i \equiv \mu_i + v_i$ is a composite error term comprising v_i , which is normally distributed with zero mean, and μ_i , which is usually assumed to be half-normally distributed and negative when the frontier is fitted as an upper envelope in the case of a profit function and positive when the frontier is fitted as a lower envelope as in the case of a cost function. β are parameters of the deterministic kernel, $F(z_i, \tau_i, \phi_i, \theta_i | \beta)$, of the stochastic frontier. The i^{th} bank's inefficiency is usually estimated by the mean of the conditional distribution of μ_i given ε_i , i.e., $E(\mu_i | \varepsilon_i)$. The difference between best-practice and achieved performance gauges managerial inefficiency in terms of either excessive cost – *cost inefficiency* – or lost profit – *profit inefficiency*. Expressing the shortfall and excess as ratios of their frontier (best-practice) values yields profit and cost inefficiency ratios. While the fitted stochastic frontier identifies best-practice performance of the banks in the sample, it cannot explain the behavior of inefficient banks. A number of papers have surveyed investigations of bank performance using these concepts: for example, Berger and Humphrey (1997), Berger and Mester (1997), and Berger (2007).

As discussed in Mester (forthcoming), since inefficiency is derived from the regression residual, selection of the characteristics of the banks and the environmental variables to include in the frontier estimation is particularly important. These variables define the peer group that determines best-practice performance against which a particular bank's performance is judged. If something extraneous to the production process is included in the specification, this might lead to too narrow a peer group and an overstatement of a bank's level of efficiency. Moreover, the variables included determine which type of inefficiency gets penalized. If bank location, e.g., urban vs. rural, is included in the frontier, then an urban bank's performance would be judged against other urban banks but not against rural banks, and a rural bank's performance would be

consuming perquisites, shirking, discriminating prejudicially, taking too much or too little risk to enhance their

judged against other rural banks. If it turned out that rural banks are more efficient than urban banks, all else equal, the inefficient choice of location would not be penalized. An alternative to including the variable in the frontier regression is to measure efficiency based on a frontier in which it is omitted and then to see how it correlates with efficiency. Several papers have looked at the correlations of efficiency measures and exogenous factors, including Mester (1993), Mester (1996), Mester (1997), and Berger and Mester (1997). Mester (1997) shows that estimates of bank cost efficiency can be biased if bank heterogeneity is ignored. See also Bos, Heid, Koetter, Kolari, and Kool (2005) on the issue of whether certain differences in the economic environment belong in the definition of the frontier.

Either the average cost function or cost frontier can be used to measure *scale economies*, which refer to how the bank's scale of operations (its size) is related to cost and give a measure of whether the bank is operating at an optimal scale. A bank is operating with scale economies if a one percent increase in scale leads to a less than one percent increase in cost; it is operating with scale diseconomies if a one percent increase in scale leads to a greater than one percent increase in costs; it is operating with constant returns to scale if a one percent increase in scale leads to a one percent increase in cost. *Scope economies* refer to whether the bank is producing the optimal combination of products to minimize cost (or maximize profits). In particular, a bank is operating with scope economies if the cost of producing the bank's product bundle is less than the cost of separating the bundle into specialized firms. The bank is operating with scope diseconomies if specialized banks could produce the product mix more cheaply.

Typically in the literature, the cost and profit functions or frontiers are measured without considering the bank's capital structure or bank's choice of risk. This is a serious omission since both are important parts of banking technology. Banks' production technologies embody their

ability to diversify and offset a variety of risks, and the production decisions managers make reflect their incentives to take on risks as well as to diversify them. Modern banking theory emphasizes managers' contrasting incentives for risk-taking. On the one hand, increased risk-taking exploits the risk-taking subsidy of explicit and implicit, mispriced deposit insurance, while, on the other hand, reduced risk-taking protects a bank from costly episodes of financial distress involving liquidity crises, regulatory intervention, and even forfeiture of the bank's valuable charter. For most banks, valuable investment opportunities make trading profitability for reduced risk a value-maximizing strategy. Reducing risk can involve not just producing assets with lower expected profit, but also incurring higher costs to manage risks.

When market-priced risk varies across production plans, the discount rate on profit will also vary across firms so that the production plan that maximizes expected profit may not maximize the discounted value of expected profit. Modeling the behavior of value-maximizing managers requires a more general objective function than profit maximization. Hughes, Lang, Mester, and Moon (1996, 1999, 2000) incorporate risk into managers' choice of production plans by defining managerial utility as a function of profit and the production plan (i.e., the choice of inputs and outputs). Technology defines all feasible production plans. The utility function ranks feasible production plans according to the utility the managers derive from each production plan. Each production plan is linked to a subjective probability distribution of profit by managers' beliefs about the probability distribution of future economic states and how these states interact with feasible production plans to determine profit. Thus, managerial utility expressed as a function of profit and the production plan is equivalent to utility expressed as a function of subjective, conditional probability distributions of profit. Hence, it allows managers to rank

production plans not just by their expected profit, the first moment of their distribution, but also by higher moments that capture the risk of production plans.

This managerial utility function is also sufficiently general that it can also account for rankings of production plans that reflect agency problems. To the extent that managers are able to pursue personal objectives that sacrifice firm value, such as empire building and risk avoidance, maximizing utility need not be the same as maximizing value, and the utility function can represent such rankings. Thus, unlike the standard maximum profit function and minimum cost function, this utility framework is able to explain *inefficient* as well as efficient managerial decisions.

To specify the performance equation (1), Hughes, Lang, Mester, and Moon (1996, 1999, 2000) adapt the Almost Ideal Demand System to derive a *utility-maximizing* profit equation and its associated input demand equations. This profit function does not necessarily maximize profit, since it follows from managers' assessment of risk and risk's effect on asset value and perhaps their job security. The profit function also might not represent value-maximizing output production plans or risk-expected return choices, to the extent that there are agency costs and managers are able to pursue non-value-maximizing objectives. Profit maximization (cost minimization) can be tested by noting that the standard translog profit (cost) function and share equations are nested within the model and can be recovered by imposing the parameter restrictions implied by profit maximization (cost minimization) on the coefficients of this adapted system. Hughes, Lang, Mester, and Moon (1996, 1999, 2000) test these restrictions and reject the hypothesis of profit maximization (and cost minimization) in their applications.

Since the utility-maximizing profit function explains inefficient as well as efficient production, it cannot be fitted as a frontier. To gauge inefficiency, Hughes, Lang, Mester, and

Moon (1996) and Hughes, Mester, and Moon (2001) estimate a best-practice risk-return frontier and measure inefficiency relative to it. The estimated utility-maximizing profit function yields a measure of expected profit for each bank in the sample, and, when divided by equity capital, the expected profit is transformed into expected return on equity, $E(\pi_i/k_i)$. Each bank's expected (or, predicted) return is a function of its production plan and other explanatory variables. When the estimation of the profit function allows for heteroscedasticity, the standard error of the predicted return (profit), σ_i , a measure of econometric prediction risk, is also a function of the production plan and other explanatory variables and varies across banks in the sample.⁵ The estimation of a stochastic frontier similar to (2) gives the highest expected return at any particular risk exposure:

$$(3) \quad E(\pi_i/k_i) = \alpha_0 + \alpha_1 \sigma_i + \alpha_2 \sigma_i^2 - \mu_i + v_i,$$

where v_i is a two-sided error term representing noise, and μ_i is a one-sided error term representing inefficiency. A bank's *return inefficiency* is the difference between its potential return and its noise-adjusted expected return, gauged among its peers with the same level of return risk. (Note, however, that if a bank's managers are taking too much or too little risk relative to the value-maximizing amount, this inappropriate level of risk is not taken into account by this measure of inefficiency.)

Koetter (2006) uses the model of managerial utility maximization and the associated measure of risk-return efficiency developed in Hughes, Lang, Mester, and Moon (1996, 1999, 2000) to investigate the efficiency of universal banks in Germany between 1993 and 2004. He

⁵ Note that the estimated profit (or return) function resembles a multi-factor model where the factors are the explanatory variables in the profit function. The regression coefficients can be interpreted as marginal returns to the explanatory variables, and the standard error of the predicted return, a function of the variance-covariance matrix of the estimated marginal returns, resembles the variance of a portfolio return. Hughes (1999) and Hughes, Mester, and Moon (2001) report that the regression of $\ln(\text{market value of equity})$ on $\ln(E(\pi_i/k_i))$ and $\ln(\sigma_i)$ for 190 publicly

compares the measure of return efficiency with cost and profit efficiency estimated by standard formulations and finds evidence that *efficient* banks using a low-risk investment strategy score poorly in terms of standard profit efficiency measures, since they also expect lower profit.

Hughes, Mester, and Moon (2001) take this a step further by recognizing that the utility-maximizing choices of bank managers need not be value maximizing to the extent that there are agency problems within the firm and managers are able to pursue their own, non-value-maximizing objectives. To identify the value-maximizing banks among the set of all banks, they select the quarter of banks in the sample that have the highest predicted return efficiency. These banks are the mostly likely group to be maximizing value or, at least, producing with the smallest agency costs. One can use this set of efficient banks to gauge characteristics of the value-maximizing production technology. For example, mean scale economies across this set of banks would indicate whether there were scale economies as banks expand output along a path that maximizes value. In contrast, mean scale economies across *all* banks would indicate whether there were scale economies as banks expand output along a path that maximizes managers' utility, but this can differ from the value-maximizing expansion path to the extent that managers are able to pursue their own objectives and these objectives differ from those of outside owners.

While the model of managerial utility maximization yields a structural utility-maximizing profit function that includes as special cases the standard maximum profit function and a value-maximizing profit function, it is, nevertheless, based on accounting measures of performance. An alternative model developed by Hughes and Moon (2003) gauges performance using the market value of assets. They develop a *utility-maximizing q-ratio function* derived from a model where managers allocate the potential (frontier) market value of their firm's assets between their

traded bank holding companies has an R-squared of 0.96, which implies that the production-based measures of expected return and risk explain a large part of a bank's market value.

consumption of agency goods (market-value inefficiency) and the production of market value, which, given their ownership stake, determines their wealth. The utility function is defined over wealth and the value of agency goods and is conditioned on capital structure, outside blockholder ownership, stock options held by insiders, and other managerial incentive variables. The authors derive a utility-maximizing demand function for market value and for agency goods (inefficiency). Hence, their q -ratio equation is *structural* and, consequently, enjoys the properties of a well-behaved consumer demand function. The authors use these properties to analyze the relationship between value (or inefficiency) and the proportion of the firm owned by insiders, which is their opportunity cost of consuming agency goods.

I.C. The nonstructural approach to bank efficiency measurement

The *nonstructural approach* to bank performance measurement usually focuses on achieved performance and measures y_i , in equation (1) by a variety of financial ratios, e.g., return-on-asset, return-on-equity, or the ratio of fixed costs to total costs. However, some applications have used measures of performance that are based on the market value of the firm (which inherently incorporates market-priced risk), e.g., Tobin's q -ratio (which is the ratio of the market value of assets to the book value of assets), the Sharpe ratio (which measures the ratio of the firm's expected excess return over the risk-free return to the volatility of this excess return (as measured by the standard deviation of the excess return)), or an event study's cumulative abnormal return, or CAR (the cumulative error terms of a model predicting banks' market return around a particular event). Other applications have measured performance by an inefficiency ratio obtained by estimating either a nonstructural or structural performance equation as a frontier. The nonstructural approach then explores the relationship of performance to various bank and environmental characteristics, including the bank's investment strategy, location,

governance structure, and corporate control environment. For example, the nonstructural approach might investigate technology by asking how performance ratios are correlated with asset acquisitions, the bank's product mix, whether the bank is organized as a mutual or stock-owned firm, and the ratio of outside to inside directors on its board. While informal and formal theories may motivate some of these investigations, no general theory of performance provides a unifying framework for these studies.

Using the frontier methods in a nonstructural approach, Hughes, Lang, Moon, and Pagano (1997) proposed a proxy for Jensen and Meckling's agency cost: a frontier of the market value of assets fitted as a potentially nonlinear function of the book-value investment in assets and the book value of assets squared. This frontier gives the highest potential value observed in the sample for any given investment in assets. For any bank, the difference between its highest potential value and its noise-adjusted achieved value represents its lost market value – a proxy for agency cost (*X*-inefficiency). Several studies have used either this systematic lost market value or the resulting noise-adjusted q-ratio to measure performance: Baele, DeJonghe, and Vennet (2006), Hughes, Lang, Mester, Moon, and Pagano (2003), DeJonghe and Vennet (2005), Hughes and Moon (2003), Hughes, Lang, Mester, and Moon (1999), and Hughes, Mester, and Moon (2001).

Habib and Ljungqvist (2005) specified an alternative market-value frontier as a function of a variety of managerial decision variables, including size, financial leverage, capital expenditures, and advertising expenditures. Thus, the peer grouping on which the frontier is estimated is considerably narrower than the wide grouping based on investment in assets, and inefficient choices of these conditioning values are not accounted for in the measurement of agency costs.

I.D. Specifying outputs and inputs in structural models of production

In estimating the standard cost or profit function or the managerial utility maximization model, one must specify the outputs and inputs of bank production. The intermediation approach focuses on the bank's production of intermediation services and the total cost of production, including both interest and operating expenses. Outputs are typically measured by the dollar volume of the bank's assets in various categories. (As mentioned above, an exception is Mester (1992), which to account for the bank's screening and monitoring activities, measured outputs as loans previously purchased, which require only monitoring, loans currently originated for the bank's own portfolio, loans currently purchased, and loans currently sold.) Inputs are typically specified as labor, physical capital, deposits and other borrowed funds, and, in some studies, equity capital. While the intermediation approach treats deposits as inputs, there has been some discussion in the literature about whether deposits should be treated as an output since banks provide transactions services for depositors. Hughes and Mester (1993) formulated an empirical test for determining whether deposits act as an input or output. Consider variable cost, VC , which is the cost of nondeposit inputs and is a function of the prices of nondeposit inputs, w , output levels, q , other variables affecting the technology, τ , and the level of deposits, x . If deposits are an input, then $\partial VC/\partial x < 0$: increasing the use of some input should decrease the expenditures on other inputs. If deposits are an output, then $\partial VC/\partial x > 0$: output can be increased only if expenditures on inputs are increased. Hughes and Mester's empirical results indicate insured and uninsured deposits are inputs at banks in all size categories.

I.E. Specifying capital structure in performance equations

As discussed above, typically, cost and profit functions are measured without considering the bank's capital structure. However, the newer literature recognizes the importance of bank

managers' choice of risk and capital structure on bank performance. Some of the first structural models to include equity capital as an input are Hancock (1985, 1986), McAllister and McManus (1993), Hughes and Mester (1993), Clark (1996), and Berger and Mester (1997).

As discussed in Hughes and Mester (1993), Hughes (1999), Mester (forthcoming), and Berger and Mester (1997), a bank's insolvency risk depends not only on the riskiness of its portfolio but on the amount of financial capital it has to absorb losses. Insolvency risk affects bank costs and profits via the risk premium the bank has to pay for uninsured debt, through the intensity of risk management activities the bank undertakes, and through the discount rate applied to future profits. A bank's capital level also directly affects costs by providing an alternative to deposits as a funding source for loans.

Most studies use the cash-flow (accounting) concept of cost, which includes the interest paid on debt (deposits) but not the required return on equity, as opposed to economic cost, which includes the cost of equity. Failure to include equity capital among the inputs can bias efficiency measurement. If a bank were to substitute debt for some of its financial equity capital, its accounting (cash-flow) costs could rise, making the less-capitalized bank appear to be more costly than a well-capitalized bank. To solve this problem, the level of equity capital can be included as a quasi-fixed input in the cost function. The resulting cost function captures the relationship of cash-flow cost to the level of equity capital, and the (negative) derivative of cost with respect to equity capital – the amount by which cash-flow cost is reduced if equity capital is increased – gives the shadow price of equity. The shadow price of equity will equal the market price when the amount of equity minimizes cost or maximizes profit. Even when the level of equity does not conform to these objectives, the shadow price nevertheless provides a measure of its opportunity cost. Hughes, Mester, and Moon (2001) find that the mean shadow price of

equity for small banks is significantly smaller than that of larger banks. This suggests that smaller banks over-utilize equity relative to its cost-minimizing value, perhaps to protect charter value. On the other hand, larger banks appear to under-utilize equity relative to its cost-minimizing value, perhaps to exploit a deposit subsidy and the subsidy due to the Too-Big-To-Fail Doctrine.

I.F. Specifying output quality in the performance equation

In measuring efficiency, one should control for differences in output quality to avoid labeling unmeasured differences in product quality as differences in efficiency. Controls for loan quality, e.g., nonperforming loans to total loans by loan category or loan losses, are sometimes included in the cost or profit frontier as controls (see Mester, forthcoming, for further discussion). As discussed in Berger and Mester (1997), whether it is appropriate to include nonperforming loans or loan losses in the cost or profit function depends on the extent to which these variables are exogenous. They would be exogenous if caused by economic shocks (bad luck), but could be endogenous to the extent that management is inefficient or has made a conscious decision to cut short-run expenses by cutting back on loan origination and monitoring resources. Berger and Mester (1997) attempt to solve this problem by using the ratio of nonperforming loans to total loans in the bank's state as a control variable. This state average would be nearly entirely exogenous to any one bank, but can control for negative shocks that affect bank output quality.

The variable, nonperforming loans, can also play a role as a quasi-fixed "input" whose quantity rather than price is included in the performance equation. As such, its "cost" is excluded from the performance metric, either cost or profit. Its price is the expected loan-loss rate. Hence, when the cost of nonperforming loans, i.e., loan losses, is excluded from the

performance measure, a case can be made for including the level of nonperforming loans, and when the performance measure is net of loan losses, the logic suggests that the loss rate be included in the specification of the performance equation.

II. Applications of the structural approach

II.A. Performance in relation to organizational form, governance, regulation, and market discipline

An increasing number of papers using structural models are exploring the importance of governance and ownership structure to the performance of banks. The structural model is first used to obtain a frontier-based measure of inefficiency. Then inefficiency is regressed on a set of explanatory variables.

Using confidential regulatory data on small, closely held commercial banks, DeYoung, Spong, and Sullivan (2001) use a stochastic frontier to measure banks' profit efficiency. They find banks that hire a manager from outside the group of controlling shareholders perform better than those with owner-managers; however, this result depends on motivating the hired managers with sufficient holdings of stock. They calculate an optimal level of managerial ownership that minimizes profit inefficiency. Higher levels of insider holdings lead to entrenchment and lower profitability.

Berger and Hannan (1998) consider the relationship of bank cost efficiency, estimated by a stochastic frontier, to product market discipline, gauged by a Herfindahl index of market power. They find that the reduced discipline of concentrated markets is associated with a loss of cost efficiency far more significant than any welfare loss due to monopoly pricing.

DeYoung, Hughes, and Moon (2001) use the model of managerial utility maximization developed by Hughes, Lang, Mester, and Moon (1996, 2000) to estimate expected return and return risk. Using these values, they estimate a stochastic risk-return frontier as in equation (3)

to obtain each bank's return inefficiency. They consider how banks' supervisory CAMEL ratings are related to their size, their risk-return choice, and their return inefficiency. They find that the risk-return choices of efficient banks are not related to their supervisory rating, while higher-risk choices of inefficient banks are penalized with poorer ratings. Moreover, the risk-return choices of large inefficient banks are held to a stricter standard than smaller banks and large efficient banks.

Two studies by Mester (1991, 1993) investigate differences in scale and scope measures for stock-owned and mutual savings and loans by estimating average cost functions. She finds evidence of agency problems at mutual S&Ls, as evidenced by diseconomies of scope, prior to the industry's deregulation, and evidence that these agency costs were lessened after the deregulation in the mid-1980s.

Using data for the period 1989-1996, Altunbas, Evans, and Molyneux (2001) estimate separate and common frontiers for three organizational forms in German banking: private commercial, public (government-owned) savings, and mutual cooperative banks. They argue that the same technology of intermediation is available to all so that the choice of technology is a management decision whose efficiency should be compared among all types of forms. The private sector appears to be less profit and cost efficient than the other two sectors. These results are especially clear in the case of the common frontier, but they are also obtained from the estimation of separate frontiers.

II.B. Uncovering evidence of scale economies by accounting for risk and capital structure

Berger and Mester (1997) use data on the almost 6000 U.S. commercial banks that were in continuous existence over the six-year period 1990-1995. They estimate scale economies, cost X-efficiency, and profit X-efficiency for banks in different size categories based on their

preferred model that incorporates asset quality, financial capital, and off-balance-sheet assets and based on several alternative specifications. In the preferred model, which includes financial capital, they find significant cost scale economies for banks in each size class: the typical bank would have to be two to three times larger in order to maximize cost scale efficiency for its product mix and input prices.

Hughes and Mester (1998) use 1989 and 1990 data on U.S. banks with assets over \$1 billion and estimate cost function conditioned on the level of financial capital. They find that banks do not hold the cost-minimizing level of capital and that the level of capitalization increases less than proportionately with assets. They find significant scale economies across banks of all size in the sample.

Hughes, Mester, and Moon (2001) apply a model of managerial utility to data on U. S. bank holding companies to consider how incorporating capital structure and endogenous risk-taking into the production model affects the ability of the empirical investigator to detect scale economies. For example, better diversification may lead to a lower cost of risk and an incentive to increase risk-taking for greater profitability. The increased risk-taking may be costly. If larger banks are better diversified and more risky than smaller banks, this source of scale economies may be hard to detect without accounting for endogenous risk-taking: the increase in cost due to the increased risk-taking can lead to the conclusion that there are no economies of scale. The authors provide evidence that better diversification is associated with larger scale economies, and increased risk-taking and inefficiency are related to smaller scale economies.

Bossone and Lee (2004) use the Hughes and Mester (1998) and Hughes, Mester, and Moon (2001) methodologies to study the relationship between productive efficiency and the size of the financial system. Using data on 875 commercial banks from 75 countries, they estimate a

cost function and measure scale economies allowing for banks' endogenous choice of risk and financial capital. Consistent with the results from Hughes and Mester (1998) and Hughes, Mester, and Moon (2001), they find significant scale economies that are increasing with the size of the financial system. They also find that small banks in larger financial systems are more cost efficient than those in small financial systems. They interpret their findings as evidence of what they call "systemic scale economies."

Berger and Mester (2003) investigate cost and profit productivity, where productivity is measured as a combination of technological change (i.e., changes in the best-practice frontier) and changes in inefficiency, holding constant the exogenous environmental variables. (This discussion is taken largely from Mester (forthcoming).) They find that during 1991-1997, cost productivity in the banking industry worsened while profit productivity improved substantially and concluded this was because revenue-based productivity changes are not accounted for in measuring cost productivity. Banks have been offering wider varieties of financial services and have been providing additional convenience, which may have raised costs but also raised revenues by more than the cost increases. They also found that banks involved in merger activity might be responsible for their main findings. The merging banks had greater cost productivity deterioration and profit productivity improvements than other banks. Merging banks may have also improved their profit performance, on average, by shifting their portfolios into investments with higher risk and higher expected return to take advantage of the diversification gains from mergers, as suggested by the work of Hughes, Lang, Mester, and Moon (1996) and Hughes, Mester, and Moon (2001).

III. Applications of the nonstructural approach

III.A. Measuring the value of investment opportunities ("charter value")

The value of a bank's investment opportunities is often measured by Tobin's q-ratio; however, in the presence of agency cost the q-ratio captures only the ability of the incumbent managers to exploit these opportunities. Ideally, the value of investment opportunities should be gauged independently of the ability and actions of the current management. Hughes, Lang, Moon, and Pagano (1997) and Hughes, Lang, Mester, Moon, and Pagano (2003) propose a measure based on fitting a stochastic frontier to the market value of assets as a function of the book value of assets and variables characterizing the market conditions faced by banks. These conditions include a Herfindahl index of market power and the macroeconomic growth rate. The fitted frontier gives the highest potential value of a bank's assets in the markets in which it operates. Thus, this potential value is conditional on the location of the bank and represents the value the bank would fetch in a competitive auction. Hughes, Lang, Moon, and Pagano (1997) define this value as the bank's "charter value" – its value in a competitive auction.

III.B. Measuring the performance of business and capital strategies

Several papers have used the nonstructural performance equation to examine the relationship between bank value and bank capital structure. Hughes, Lang, Moon, and Pagano (1997) regress performance measured by Tobin's q-ratio and market-value inefficiency on a number of variables characterizing bank production. Calomiris and Nissim (2007) regress the ratio of the market value of equity to its book value on a similar list of variables. De Jonghe and Vennet (2005) apply the market-value frontier of Hughes, Lang, Moon, and Pagano (1997) to derive a noise-adjusted measure of Tobin's q, which they use to evaluate how leverage and market power are related to value. All three studies find evidence that banks follow dichotomous strategies for enhancing value as predicted by Marcus (1984): a lower risk, lower leverage strategy and a higher risk, higher leverage strategy.

III.C. Relationship of ownership structure to bank value

Jensen and Meckling (1976) defined agency cost as the difference in value of a firm owned entirely by its manager (so that there are no agency problems) and one where the manager does not own all of the firm. Since firms with no agency costs should out-perform those with agency problems, some studies have sought evidence of agency costs by looking for a correlation between firm value measured by Tobin's q-ratio and variables characterizing potential agency problems, such as the proportion of the firm owned by managers and the proportion owned by outside blockholders.

In an influential study, Morck, Shleifer, and Vishny (1988) hypothesized that managerial ownership creates two contrasting incentives: a higher ownership stake, first, better aligns the interests of managers and outside owners and, second, enhances managers' control over the firm and makes it harder for managers to be ousted when they are not efficient. Measuring performance by Tobin's q, these authors provide evidence that the so-called alignment-of-interests effect dominates the entrenchment effect at lower levels of managerial ownership, while the entrenchment effect dominates over a range of higher levels.

Studies that attempt to measure the *net* effect of the alignment and entrenchment effects on firm valuation cannot identify these effects individually – only their sum in the form of the sign of a regression coefficient or a derivative of a regression equation. Adams and Santos (2006) cleverly isolate the entrenchment effect by considering how the proportion of a bank's common stock controlled but not owned by the bank's own trust department is statistically related to the bank's economic performance. The voting rights exercised by management through the trust department enhance management's control over the bank but do not align their

interests with outside shareholders', since the beneficiaries of the trusts, not the managers, receive the dividends and capital gains and losses.

Caprio, Laeven, and Levine (2003) study the effect of ownership, shareholder protection laws, and supervisory and regulatory policies on the valuations of banks around the world. The authors construct a database of 244 banks – in each of 44 countries. They measure performance by Tobin's q-ratio and by the ratio of the market value of equity to the book value of equity. They find evidence that banks in countries with better protection of minority shareholders are more highly valued; bank regulations and supervision have no significant effect on bank value; the degree of cash-flow rights of the largest owner has a significant positive effect on bank value; and an increase in ownership concentration has a larger positive effect on valuation when the legal protection of minority shareholders is weak.

III.D. Relationship of mergers and takeovers to bank value

Brook, Hendershott, and Lee (1998) examine the stock market reaction to the passage of the Interstate Banking and Branching Efficiency Act (IBBEA) of 1994. They find significantly positive abnormal returns that are negatively related to a bank's prior performance. Apparently, the increased probability of a takeover following the passage of IBBEA improves the value of underperforming banks more than better performing banks. This increase in value is offset among banks whose managers show evidence of entrenchment, such as higher insider ownership, lower outside blockholder ownership, and less independent boards.

If the threat of a takeover disciplines managers and improves profitability, differences in takeover restrictions across states imply differences in the threat of a takeover. Schranz (1993) finds that banks in states with a more active takeover market are more profitable than banks in states restricting takeover activity.

Hughes, Lang, Mester, Moon, and Pagano (2003) examine U.S. bank holding companies and find evidence of managerial entrenchment among banks with higher levels of insider ownership, more valuable growth opportunities, poorer financial performance, and smaller asset size. When managers are not entrenched, asset acquisitions and sales are associated with reduced market value inefficiency. When managers are entrenched, sales are associated with smaller reductions in inefficiency, while acquisitions are associated with greater inefficiency.

DeLong (2001) studies 280 domestic U.S. bank mergers from 1988 through 1995. Gauging performance by the CARs of the mergers, she finds that mergers that focus activity and geography increase shareholder value, while diversifying mergers do not.

IV. Conclusions

Great strides have been made in the theory of bank technology in terms of explaining banks' comparative advantage in producing informationally intensive assets and financial services and in diversifying or offsetting a variety of risks. Great strides have also been made in explaining sub-par managerial performance in terms of agency theory and in applying these theories to analyze the particular environment of banking. In recent years, the empirical modeling of bank technology and the measurement of bank performance have begun to incorporate these theoretical developments and yield interesting insights that reflect the unique nature and role of banking in modern economies.

Bibliography

Adams, R.B. and Santos, J.A.C. (2006). Identifying the effect of managerial control on firm performance, *Journal of Accounting and Economics* **41**, 55-85.

Aghion, P., Alesina, A. and Trebbi, F. (May 2007). Democracy, technology, and growth, Working Paper, Department of Economics, Harvard University.

Altunbas, Y., Evans, L. and Molyneux, P. (2001). Bank ownership and efficiency, *Journal of Money, Credit, and Banking* **33**, 926-954.

Baele, L., DeJonghe, O. and Vander Venet, R. (August 2006). Does the stock market value bank diversification? Working Paper No. 2006/402, Department of Financial Economics, Ghent University.

Berlin, M. and Mester, L.J. (1999). Deposits and relationship lending, *Review of Financial Studies* **12**, 579-607.

Berger, A.N. (2007). International comparisons of banking efficiency, *Financial Markets, Institutions and Instruments* **16**, 119–144.

Berger, A.N. and Hannan, T.H. (1998). The efficiency cost of market power in the banking industry: a test of the ‘quiet life’ and related hypotheses, *Review of Economics and Statistics* **80**, 454-465.

Berger, A.N. and Humphrey, D. B. (1997). Efficiency of financial institutions: international survey and directions for future research, *European Journal of Operational Research* **98**, 175-212.

Berger, A.N. and Mester, L.J. (1997). Inside the black box: what explains differences in the efficiencies of financial institutions, *Journal of Banking and Finance* **21**, 895-947.

Berger, A.N. and Mester, L.J. (2003). Explaining the dramatic changes in performance of U.S. banks: technical change, deregulation, and dynamic changes in competition, *Journal of Financial Intermediation* **12**, 57-95.

Bhattacharya, S. and Thakor, A. (1993). Contemporary banking theory, *Journal of Financial Intermediation* **3**, 2-50.

Bos, J.W.B., Heid F., Koetter, M., Kolari, J.W. and Kool, C.J.M. (2005). Inefficient or just different? Effects of heterogeneity on bank efficiency scores, Deutsche Bundesbank Discussion Paper No. 2.

Bossone, B. and Lee, J.-K. (2004). “In finance, size matters: the ‘systemic scale economies’ hypothesis, IMF Staff Papers, 51:1.

Brook, Y., Hendershott, R. and Lee, D. (1998). The gains from takeover deregulation: evidence from the end of interstate banking restrictions, *Journal of Finance* **53**, 2185-2204.

Calomiris, C.W. and Kahn, C. M. (1991). The role of demandable debt in structuring optimal banking arrangements, *American Economic Review* **70**, 312-326.

Calomiris, C.W. and Nissim, D. (2007). Activity-based valuation of bank holding companies, Working Paper 12918, National Bureau of Economic Research.

Caprio, G., Laeven, L. and Levine, R. (2003). Governance and bank valuation, Working Paper 10158, National Bureau of Economic Research.

Clark, J. (1996). Economic cost, scale efficiency and competitive viability in banking, *Journal of Money, Credit, and Banking* **28**, 342-364.

DeJonghe, O. and Vander Vennet, R. (2005). Competition versus agency costs: an analysis of charter values in European banking, Working Paper, Ghent University.

DeLong, G.L. (2001). Stockholder gains from focusing versus diversifying bank mergers, *Journal of Financial Economics* **59**, 221-252.

Demirgüç-Kunt, A., Kane, E.J. and Laeven, L. (January 2007). Determinants of deposit-insurance adoption and design, NBER Working Paper No. 12862.

DeYoung, R.E., Hughes, J.P. and Moon, C.-G. (2001). Efficient risk-taking and regulatory covenant enforcement in a deregulated banking industry, *Journal of Economics and Business* **53**, 255-282.

DeYoung, R., Spong, K. and Sullivan, R.J. (2001). Who's minding the store? Motivating and monitoring hired managers at small, closely held commercial banks, *Journal of Banking and Finance* **25**, 1209-1243.

Fiordelisi, Franco (2007). Shareholder value efficiency in European banking, *Journal of Banking and Finance* **31**, 2151-2171.

Flannery, M.J. (1994). Debt maturity and the deadweight cost of leverage: optimally financing banking firms, *American Economic Review* **84**, 320-331.

Habib, M.A. and Ljungqvist, A. (2005). Firm value and managerial incentives: a stochastic frontier approach, *Journal of Business* **78**, 2053-2093.

Hancock, D. (1985). The financial firm: production with monetary and nonmonetary goods, *Journal of Political Economy* **93**, 859-880.

Hancock, D. (1986). A model of the financial firm with imperfect asset and deposit liabilities, *Journal of Banking and Finance* **10**, 37-54.

Hughes, J.P. (1999). Incorporating risk into the analysis of production, presidential address to the Atlantic Economic Society, *Atlantic Economic Journal* **27**, 1-23.

Hughes, J.P., Lang, W., Mester, L.J. and Moon C.-G. (2000). Recovering risky technologies using the almost ideal demand system: an application to U.S. banking, *Journal of Financial Services Research* **18**, 5-27.

Hughes, J.P., Lang, W., Mester, L.J. and Moon C.-G. (1999). The dollars and sense of bank consolidation, *Journal of Banking and Finance* **23**, 291-324.

Hughes, J.P., Lang, W., Mester, L.J. and Moon C.-G. (1996). Efficient banking under interstate branching, *Journal of Money, Credit, and Banking* **28**, 1045-1071.

Hughes, J.P., Lang, W., Mester, L.J., Moon C.-G. and Pagano, M. (2003). Do bankers sacrifice value to build empires? Managerial incentives, industry consolidation, and financial performance, *Journal of Banking and Finance* **27**, 417-447.

Hughes, J.P., Lang, W., Moon C.-G. and Pagano, M. (1997). Measuring the efficiency of capital allocation in commercial banking, Working Paper 98-2, Federal Reserve Bank of Philadelphia (revised as Working Paper 2004-1, Rutgers University Economics Department).

Hughes, J.P. and Mester, L.J. (1998). Bank capitalization and cost: evidence of scale economies in risk management and signaling, *Review of Economics and Statistics* **80**, 314-325.

Hughes, J.P. and Mester, L.J. (1993). A quality and risk-adjusted cost function for banks: evidence on the 'too-big-to-fail' doctrine, *Journal of Productivity Analysis* **4**, 293-315.

Hughes, J.P., Mester, L.J. and Moon C.-G. (2001). Are scale economies in banking elusive or illusive? Evidence obtained by incorporating capital structure and risk-taking into models of bank production, *Journal of Banking and Finance* **25**, 2169-2208.

Hughes, J.P. and Moon C.-G. (2003). Estimating managers' utility-maximizing demand for agency goods, Working Paper 2003-24, Department of Economics, Rutgers University.

Humphrey, D.B. and L.B. Pulley (1997). Banks' responses to deregulation: profits, technology, and efficiency, *Journal of Money, Credit, and Banking* **29**, 73-93.

Jensen, M.C. and Meckling, W.H. (1976). Theory of the firm: managerial behavior, agency costs and ownership structure, *Journal of Financial Economics* **5**, 305-360.

Jensen, M.C. and Meckling, W.H. (1979). Rights and production functions: an application to labor-managed firms and codetermination, *Journal of Business* **52**, 469-506.

Koetter, M. (2006). The stability of efficiency rankings when risk-preferences and objectives are different, Discussion Paper 08/2006, Series 2: Banking and Financial Studies, Deutsche Bundesbank.

- La Porta, R., Lopez-de-Silanes, F. and Shleifer, A. (2002). Government ownership of banks, *Journal of Finance* **57**, 265-301.
- Leibenstein, H. (1966). Allocative efficiency vs. 'X-efficiency,' *American Economic Review* **56**, 392-415.
- Marcus, A.J. (1984). Deregulation and bank financial policy, *Journal of Banking and Finance* **8**, 557-565.
- McAllister, P.H. and McManus, D. (1993). Resolving the scale efficiency puzzle in banking, *Journal of Banking and Finance* **17**, 389-406.
- Mester, L.J. (forthcoming). Optimal industrial structure in banking, in *Handbook of Financial Intermediation*, Boot, A. and Thakor, A. (eds.) Amsterdam: North-Holland/Elsevier.
- Mester, L.J. (First and Second Quarters 2007). Some thoughts on the evolution of the banking system and the process of financial intermediation, Federal Reserve Bank of Atlanta *Economic Review*, 67-75.
- Mester, L.J. (1997). Measuring efficiency at U.S. banks: accounting for heterogeneity is important, *European Journal of Operational Research* **98**, 230-242.
- Mester, L.J. (1996). A study of bank efficiency taking into account risk-preferences, *Journal of Banking and Finance* **20**, 1025-1045.
- Mester, L.J. (1993). Efficiency in the savings and loan industry, *Journal of Banking and Finance*, **17**, 267-286.
- Mester, L.J. (1992). Traditional and nontraditional banking: an information-theoretic approach, *Journal of Banking and Finance* **16**, 545-566.
- Mester, L.J. (1991). Agency costs among savings and loans, *Journal of Financial Intermediation* **1**, 257-278.
- Mester, L.J., Nakamura, L.I. and Renault, M. (2007). Transactions accounts and loan monitoring, *Review of Financial Studies* **20**, 529-556.
- Morck, R., Shleifer, A. and Vishny, R.W. (1988). Management ownership and market valuation: an empirical analysis, *Journal of Financial Economics* **20**, 293-315.
- Schranz, M.S. (1993). Takeovers improve firm performance: evidence from the banking industry, *Journal of Political Economy* **101**, 299-326.