2 TECHNOLOGY AND FINANCIAL INTERMEDIATION IN MULTIPRODUCT BANKING FIRMS

An Econometric Study of U.S. Banks, 1979–1982

Colin Lawrence
Robert P. Shay

INTRODUCTION

This investigation attempts measurement of the impact of computer technology upon economies of scale and scope (cost complementarities) and the elasticities of substitution between labor, financial capital, and computers. It falls firmly within the emerging literature that since 1982 has corrected earlier attempts to measure scale economies by using Cobb-Douglas production technologies. These latter functions are flawed by the assumption that each product "output" of a financial intermediary is independent of other outputs in relation to cost and that such a specification precludes U-shaped cost curves. Beginning in 1982 the use of the translog functional form allowed researchers to relax the assumption of independent outputs and to test for the existence of economies of scope. The use of this production function has also permitted estimation of U-shaped cost curves, varying elasticities of substitution in inputs across banks, as well as specific measures of scale economies for each output. R. Alton Gilbert (1984) has traced the development of the research on bank costs through six stages, the latter stage marking the begin-

The authors are indebted to Allen Berger, David Humphrey, Frank Lichtenberg, John Murray, Mark Flannery, and Michael Smirlock for critical comments on earlier drafts. We are also grateful to Julian Silk for research assistance.

ning of tests for cost complementarity and the consequent measurement of economies of scope.

Despite the improvement in methodology and techniques of measuring economies of scale, the basic finding that scale economies are found only among small banks (below $50 million in deposits) has been supported consistently by studies preceding this one. A newer related finding is that *diseconomies* of scale are found in banks where deposits total more than $50 million in deposits. One finds this difficult to believe when there is such a rush toward interstate banking and one observes the massive reliance upon computer technology and electronic banking. Can it be the omission of data from the largest banks that are at the forefront of the application of technology that is responsible for biasing the estimates of scale economies? This is doubtful, since computer applications have been adopted widely throughout the banking system. The conclusion must then be that computer technology improves efficiency by expansion in plant size due to the presence of scope economies alone. The three most recent studies have identified the existence of cost complementarities (jointness), suggesting economies of scope, although one of the studies questions the use of the translog function to test for the existence of cost complementarities (Murray and White 1983; Gilligan, Smirlock, and Marshall 1984; Benston et al. 1983).

The twin questions of the existence of scale and scope economies—subadditivity—are crucial in order to anticipate the possible effects of interstate banking and the lines of business that banks are permitted to enter. In particular, the abrogation of several clauses of the Glass-Steagall Act contained in the current deregulatory policies would imply that the newly emerging financial structure would depend critically on the existence of scale and scope economies. If the latter economies do not exist beyond relatively small bank asset sizes, there is little reason to believe that permitting large banks to expand outside of state boundaries will result in only a few giant banks with thousands of branches to service our nation.[1] If scale and/or scope economies are present, on the other hand, the possible benefits to the banking public through lower prices and a wider range of products at banks could well be appreciable. The question of market power and monopoly pricing would have to be considered if the elimination of competitors were sufficiently large to pose problems of size or restrictive bank practices among remaining banks.

Clearly, the computer and telecommunications technology is changing the face of wholesale and retail banking. On the corporate side, fund flows have been greatly speeded up by electronic cash management systems, and electronic fund transfers are replacing paper check clearing, particularly on large transfers. On the retail side, automatic teller machines (ATMs) have made it possible for retail cash dispensing, deposit, and revolving credit facilities to operate twenty-four hours daily, and home banking from on-line personal computers will make it possible for a family to conduct banking business as well as financial investment without leaving home. Point-of-sale (POS) funds transfers from banking accounts in retail stores will become commonplace as debit cards begin to gain acceptability as did credit cards. It is for these reasons that we believe technology must be having a dramatic impact on costs in banking and other financial intermediaries and will continue to do so.

So we begin our study with two fundamental objectives: First, we improve upon the methodology used by other researchers in estimating scale and scope economies by explicitly introducing computer capital into the analysis. We then estimate elasticities of substitution between labor, computers, and financial capital used to produce four banking outputs: total deposits, loans, investments, and nonbanking expenses (the provision of fee-based services such as data processing, trusts, and safety boxes). These are estimated across bank size over a four-year period, 1979–82, using the Federal Reserve's Functional Cost Data with about 650 banks each year.

Our second objective is to develop and use measures that will indicate the impact of technological developments affecting banking upon economies of scale and scope. These include automatic teller machines, point-of-scale merchandising transactions and electronic check-clearing volume.

We are presenting some initial results from our research that reveal some novel implications and problems connected with studies in bank costs. Because of the relation of the topic to current regulatory and legislative developments, we stress that our results are preliminary at this stage and should be interpreted with caution.

In contrast to other studies, we find significant economies of scale even for large banks (up to $2.5 billion in deposits). The reason that earlier studies fail to find this is that they have either ignored the multiproduct nature of the banking firm or statistically aggregated banks of all sizes in their samples. We find that the technological parameters differ across bank size. Although in some cases we do find an absence of economies of scale, there is certainly no existence of diseconomies of scale.

We also find significant economies of scope between deposits and loans as well as between deposits and investment.[2] There are also statistically significant economies of scope between fee-based banking services and deposits/investments. There appears to be significant diseconomies of scope between investment and loans. Finally, we find relatively high significant elasticities of substitution between computers and labor, often above 2 in some of the subsamples.

Our preliminary findings thus point to the likelihood that the computer/labor ratio will continue to rise dramatically as computer costs fall (relative to labor) and financial intermediaries will move toward concentrated supermarkets offering an array of financial services, thus exploiting both scope and scale economies.

THE PRODUCTION FUNCTION OF BANKS

Financial Intermediation

Theoretical models of the behavior of financial intermediaries (FIs) have traditionally focused on financial portfolio choices. This is summarized ex post by the balance sheet. In sharp contrast, empirical studies concerned with estimating scale economies in banking place emphasis on the transformation of physical inputs (labor, materials, and physical capital) into higher valued financial output (demand deposits, value of earning assets, etc.). The latter empirical studies were initiated by Benston (1965) and Bell–Murphy (1967) using a Cobb–Douglas production function. This led to a proliferation of studies on scale economies in FIs along the same methodological lines.[3]

The bulk of the empirical literature has ignored the portfolio choice aspect of FIs. For example, interest costs account for around 70 percent of total costs and yet have received very little attention.[4]

It is our contention that empirical work must synthesize financial portfolio choice and production (cost) functions for an adequate treatment of economies of scale and scope. Ignoring the interest rate, where different shareholders have varying degrees of risk aversion, could lead to serious specification error and inconsistent empirical estimation.

Neglect of the balance sheet has developed because researchers have pragmatically tended to ignore the complex multiple input-output structure of banking.

Banks do not behave like nonfinancial corporations. They have a unique or atypical production function which lies at the heart of intermediation. As Sealey and Lindley (1977) have argued, financial capital, which includes deposits, is an input in a fixed proportion production function. This means that in the absence of fee-based services including trusts, brokerage, and data services, a bank cannot expand output (earning assets) without increasing financial capital (i.e., deposits, other liabilities, equity). To avoid inconsistent estimation, this balance sheet restriction must be imposed on the cost structure.

The Dual Role of Transaction Deposits

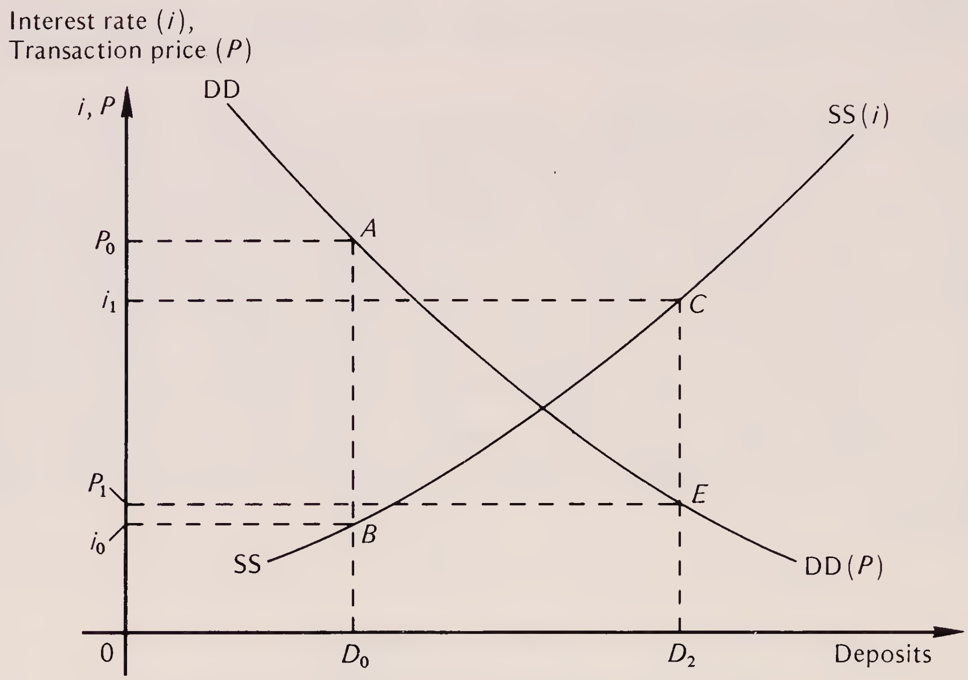
Empirical work to date has included deposits as an output provided by banks to their customers. Deposits play this role by providing liquidity or transaction services. However, there is a second role. Deposits act as an input in the production of earning assets. Although it is true that bank customers pay a fee in the form of commissions, checking fees, and so on, in return for the transaction services, they are also paid interest on their deposits for supplying financial capital to the bank.

Since deposits have a dual role, the net price paid by the FI to the customer could be either negative or positive to settle the reciprocal trade. The terms of trade will depend upon the net marginal benefits customers derive from utilizing this transaction technology relative to their rate of time preference. In the following section we derive a net supply curve of financial capital.

A Simple Exposition of the Production Function

The flow of funds for an FI are postulated as follows: D is the net inflow of homogenous deposits and L is the flow of homogenous

Figure 2-1. Customers' Demand and Supply Schedules for Deposits.

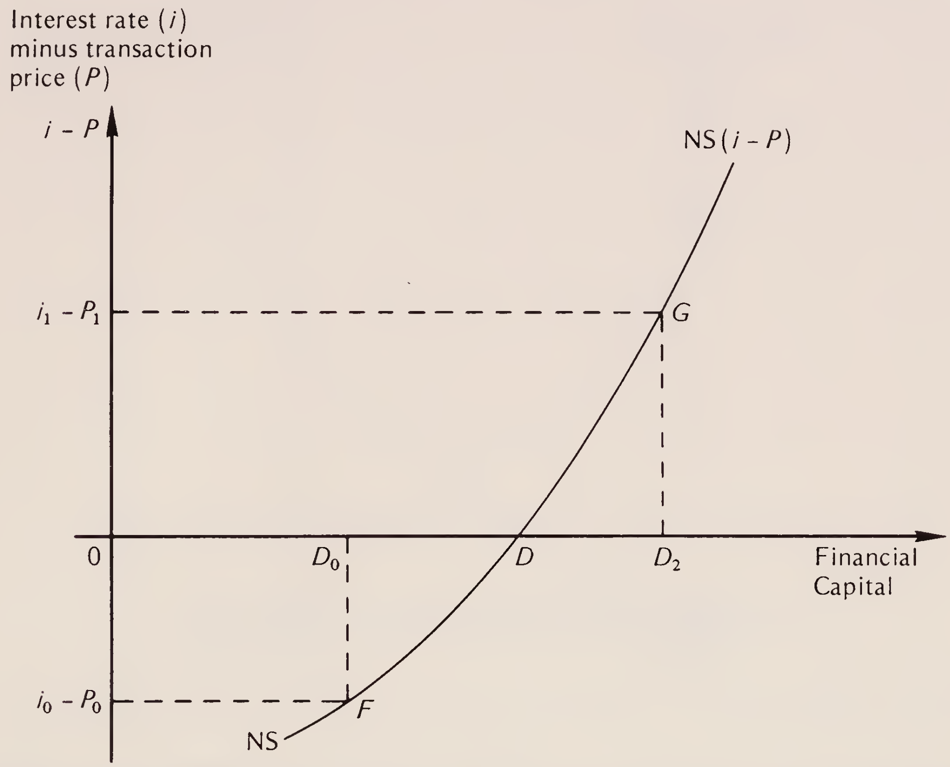


loans. The constraint applicable to FIs, that loans be funded by deposits, is also applicable to flows of funds in the analysis to follow.

For simplicity we ignore capital financing and reserve requirements. The first role is the output role—FI sells a transaction service to its customers at a rate of $\$P$ per dollar of deposits per unit of time. Customers' demand function for this service is the usual downward sloping demand curve depicted in Figure 2-1 as DD. On the other hand, although customers obtain a transaction service, they are providing the bank with a flow of financial capital. Thus agents are refraining from consumption and must be compensated. They are providing a supply of financial capital at $\$i$ per dollar/deposit per unit of time (the interest rate). The supply schedule is depicted in Figure 2-1 as the SS schedule.

The net supply schedule of financial capital provided by borrowers is derived by subtracting the price per dollar/deposit per unit of time from the interest rate per dollar/deposit at each level of deposit on the horizontal axis of Figure 2-1. This net supply curve is depicted as the NS curve in Figure 2-2. The net price $i - P$ could be either

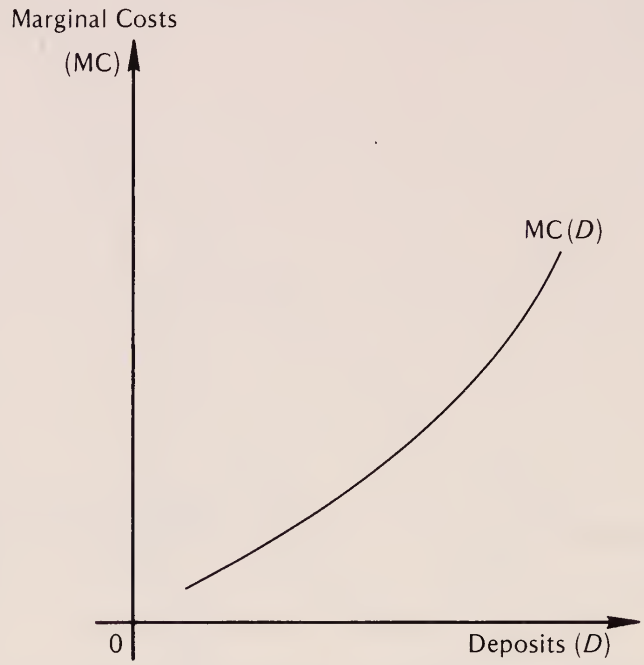
Figure 2-2. Customers' Net Supply Schedule of Financial Capital.



negative or positive. For example, at point F on NS, customers provide intermediaries with $\$D_0$ of deposits at a negative price $i_0 - P_0$; that is, they are willing to pay $\$P_0$ per dollar deposit for transaction services (point A on the DD curve in Figure 2-1) and demand $\$i_0$ per dollar deposit (point B on the SS curve in Figure 2-1). On the other hand, points above D in Figure 2-2, such as G, on the NS schedule imply a positive cost of capital, $i - P > 0$. At G, they will pay P_1 per dollar transaction services (point E on the DD curve) and demand i_1 per dollar deposit as suppliers of financial capital (point C).

An equilibrium will occur at whatever point the demand for the loan services offered by banks (not shown) intersects the net supply schedule for financial capital in the lower panel. The intersection of DD and SS represents the breakeven point where the marginal benefits of the transaction service offered to the customer equals the marginal cost of financial capital to the lender and is simply one point on the NS curve (point D).

Figure 2-3. The Supply of Transaction Deposits.

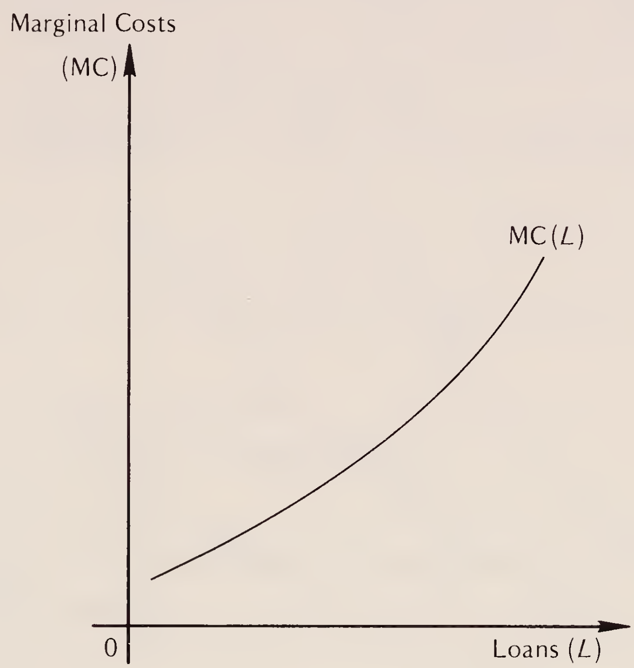


The Production Function. The FI produces transaction deposits with the aid of physical factors of production, labor, capital (buildings, computers), land, materials, and so on. For simplicity, there is only one composite factor, k, with a unit cost of $\$W$. The marginal cost schedule for supplying transaction output is shown in Figure 2-3. Similarly, the processing of loans involves physical inputs and its supply schedule is depicted in Figure 2-4. Both these schedules are determined jointly due to inseparable production functions involving transray economies of scale. The MC function for loans is not a conventional function, for it assumes that as the level of loans is expanded so is the level of financial capital. The production function for loans is described as

$$L = \min[D, g(k)]$$

where $g(k)$ is the physical production function linking the volume of loans to the value of physical inputs. Thus loans are produced with both deposits (financial capital) and physical inputs. The actual marginal cost of producing one dollar of loans is equal to the addition of the MC(D) and MC(L) schedules to the NSD$(i - P)$ shown in Figure 2-5. This schedule has been derived under the condition of the criti-

Figure 2-4. The Supply of Loans.



cal financial intermediation balance sheet restriction, namely, that $D = L$, in our simple exposition. If the FI is a price taker in the loan market, it will face a perfectly elastic supply curve at the prevailing market rate, r, as indicated in Figure 2-5. Equilibrium is established at point E, and $DD^* = L^*$ is the optimal level of the balance sheet. The optimal condition for equilibrium is that

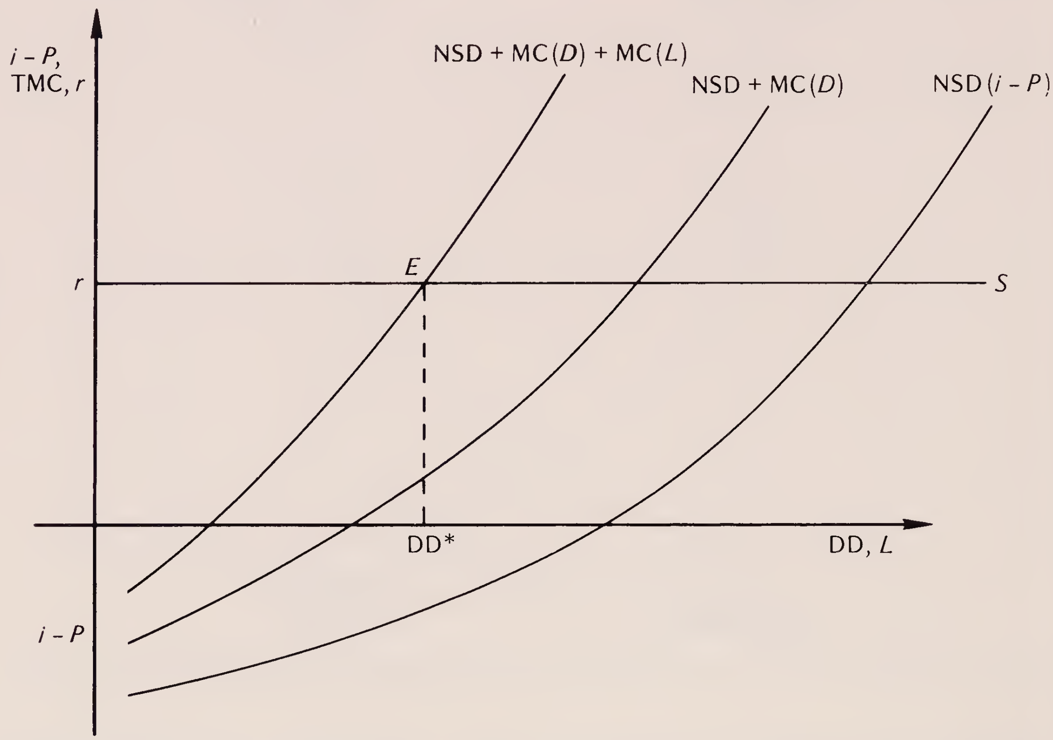
$$r \geqq MC(L) + MC(D) + (i - P) \quad .$$

The optimal level of deposits and loans for any financial intermediary depends on the yield on loans, costs of inputs, customer valuation of transaction services, and rate of time preference. Shifts of any of these components will shift the equilibrium point E.

The above exposition can be easily extended to include multiple earning assets and other banking output activities such as leasing, deposit boxes, and trusts.

An important implication of the above theory is that unless banks expand into nonbanking activity such as brokerage or trusts, then the elasticity of substitution between financial capital and computers or labor will be zero for a given quality of services.[5]

Figure 2-5. Equilibrium for the Financial Intermediary.



ECONOMETRIC METHODOLOGY

We estimate a flexible cost function for a multiproduct banking firm. We posit a priori restrictions so as to establish consistency between a well-behaved production function and the dual cost system. Moreover, following Christianson, Jorgensen, and Lau (1973), we utilize Shephard's Lemma and estimate the translog cost function together with the share equations utilizing Zellner's seemingly unrelated equations (SUR). This system is then estimated together with the cross-equation restrictions implied by Shephard's lemma and linear homogeneity in prices.[6] From the above system we can derive estimates of economies of scope and multiproduct scale economies. In addition, we can estimate specific scale economies as well as elasticities of substitution between the inputs in the model. The above procedure is carried out for each quartile in our four annual samples.[7]

The system to be estimated is as follows:

$$\text{LTC} = \alpha_0 + \sum_i^4 \alpha_i \text{LX}_i + \sum_i^3 \beta_i \text{LP}_i + \tfrac{1}{2} \sum_i^4 \sum_j^4 \alpha_{ij} \text{LX}_i \text{LX}_j + \sum_i^4 \sum_j^3 \gamma_{ij} \text{LX}_i \text{LP}_j$$

$$+ \tfrac{1}{2} \sum_i^3 \sum_j^3 \beta_{ij} \text{LP}_i \text{LP}_j + \delta_1 \text{TNO} + \delta_2 \text{ALNS} + \delta_3 \text{ADEPS} + V \qquad (2\text{-}1)$$

$$S_i = \beta_i + \sum_{j=1}^{3} \beta_{ij} \text{LP}_j + \sum_{j=1}^{4} \gamma_{ij} \text{LX}_j + U_i \qquad i = 1, 2 \qquad (2\text{-}2)\ (2\text{-}3)$$

where:

LTC	is the log of total costs
LP_1	log of interest on available funds
LP_2	log of wages
LP_3	log of computer rental
LX_1	log of deposits
LX_2	log of loans
LX_3	log of investments
LX_4	log of fee-based banking activity
TNO	number of banking offices
ALNS	log of average loan size
ADEPS	log of average deposits
S_1	share of interest in total costs
S_2	share of wages in total costs
V, U_1, U_2	random disturbances with covariance matrix Σ.

It is evident that the share equations involve no loss in degrees of freedom. Only two share equations are estimated since adding a third would lead to singularity in the error covariance matrix.[7]

The main justification for specification of the translog equation (2.1) is that deposits appear as an output—transaction services—whereas the interest on available funds appears as an input—price—the rental cost of financial capital. This is the rationale for including both deposits and interest rates in this dual specification. We have also added three extra terms to control for number of branches, average loan size, and average deposit size.[8] The above specification also includes a computer rental variable LP_3.

Economies of Scale

In the multiproduct banking firm economies of scale are estimated utilizing Baumol's ray average costs. Intuitively, we ask the question of whether or not an equiproportionate increase in all outputs (LX_1, LX_2, LX_3, and LX_4) would lead to a less than proportionate increase in total costs. If such is the case, we have declining ray average costs which imply economies of scale. Ray average costs of the multiproduct bank are defined as

$$RAC(X) = \frac{TC(X)}{\Sigma X_i} = \frac{TC(kX^\circ)}{k} \tag{2-4}$$

where $kX^\circ = X$ and $\Sigma X_1^\circ = 1$. This is the standard average cost of the composite commodity whose unit is the vector $X^\circ = [X_1^\circ, X_2^\circ, X_3^\circ, X_4^\circ]$. To test for declining ray average costs (with respect to composite output), substitute $X_i = kX_1$ for $i = 2, 3, 4$ into the translog equation and differentiate with respect to X_1. The solution is,

$$\frac{\partial LTC}{\partial LX} = \sum_{i=1}^{4} \alpha_i + \sum_{i=1}^{4}\sum_{j=1}^{4} \alpha_{ij} LX_j + \sum_{i=1}^{3}\sum_{j=1}^{4} \gamma_{ij} LP_i \tag{2-5}$$

$$\text{if } \frac{\partial LTC}{\partial LX} \leqq 1, \text{ economies of scale exist.}$$

Moreover, specific economies of scale or elasticities of costs with respect to outputs can be computed by differentiating (2-1) with respect to LX_j for $j = 1, 4$. This yields

$$\frac{\partial LTC}{\partial LX_j} = \alpha_j + \sum_{i=1}^{4} \alpha_{ij} LX_i + \sum_{i=1}^{3} \gamma_{ij} LP_i \quad \text{for } j = 1, 2, 3, 4. \tag{2-6}$$

One can observe that

$$\frac{\partial LTC}{\partial LX} = \sum_{j=1}^{4} \frac{\partial LTC}{\partial LX_j} \tag{2-7}$$

so that the sum of specific scale economies must equal the overall scale economies defined by declining ray average costs. In the estimation procedure, we calculate $\partial LTC/\partial LX$ at the mean of each quartile, where each quartile is estimated independently each year.[9]

Economies of Scope

Economies of scope or jointness in production is said to exist between any two outputs i and j if $\partial TC/\partial X_i \partial X_j < 0$. This is a sufficient condition to ensure that $TC(0, X_j) + TC(X_i, 0) \geqq TC(X_i, X_j)$.[10] Denny and Pinto (1978) have shown that the sufficient condition for cost complementary is that

$$\frac{\partial^2 LTC}{\partial LX_i LX_j} + \left(\frac{\partial LTC}{\partial X_i}\right)\left(\frac{\partial LTC}{\partial X_j}\right) < 0 \quad . \qquad (2\text{-}8)$$

In terms of translog specification (2-1), the above condition can be approximated around the mean of each variable to

$$\alpha_i \alpha_j + \alpha_{ij} < 0 \qquad \text{for all } i, j \quad i \neq j \quad . \qquad (2\text{-}9)$$

If (2-9) holds, economies of scope are said to exist. We estimate (2-9) for each pair of outputs.

Elasticities of Substitution and Derived Demand Elasticities

Allen's partial elasticities of substitution are calculated as follows,

$$\sigma_{ij} = [\beta_{ij} + S_i S_j] / S_i S_j \qquad \text{for } i \neq j$$

and

$$\Omega_j = [\beta_{jj} + S_j (S_j - 1)] / S_j$$

where σ_{ij} is the elasticity of substitution between i and j and Ω_j is the derived demand elasticity for input j with respect to price j. It should be noted that a necessary condition for the Hessian matrices' eigenvalues to be nonnegative (to ensure concavity of the cost function) is that $\Omega_j < 1$ for all j.

Other Tests

Further restrictions on the translog technology in (2-1) can be imposed to test for homotheticity, homogeneity, and separability.

The homogeneity postulate is satisfied if $\Sigma \beta_i = 1$, $\sum_i \beta_{ij} = \sum_j \beta_{ij} = 0$ and $\sum_i \gamma_{ij} = 0$. We estimate the system (2-1), (2-2), and (2-3) imposing the above homogeniety system. Equation (2-1) can also be tested for a Cobb-Douglas or constant elasticity of substitution (CES) specification. In the Cobb-Douglas case, all the second-order terms should be zero. A weaker case is that of homotheticity where each α_{ij} and γ_{ij} should be zero. F-tests can be performed to test for these restrictions.

We also tested the null hypothesis that technological and cost parameters differed across quartiles. For this purpose the unrestricted sum of squared residuals (allowing parameters to differ across quartiles) was compared to the sum of squared errors of restricted system using an F-statistic. This was calculated separately for each year, 1979-1982.

Finally, our theory of financial intermediation expressed earlier suggests that interest costs should be included in total costs and that an interest yield should be included in the cost function as a price of inputs. Many researchers have not pursued this specification. We thus estimated the system without interest costs and an interest rate and compared this set of regressions to the original regressions.

We now turn to a description of our data and our preliminary empirical findings.

THE DATA AND THEIR MEASUREMENT

Data used in this study are from the Federal Reserve Functional Cost Analysis (FCA) program from the years 1979 through 1982. After preliminary screening, 623 banks were included in the sample ranging in deposit size from a minimum of about \$6 million to a maximum of \$2.6 billion in 1982. The mean deposit size was \$141 million. The major advantage of this sample is that it presents cost estimates that are based on a standardized procedure for measurement; its major disadvantage is that it is biased downward by size and does not include any of the nation's largest banks. Another disadvantage is that different banks may participate in the FCA program each year and the individual banks cannot be identified to track a year-to-year common bank sample. For this reason we divided the banks into quartiles or deciles in order to measure economies of scale and scope by bank size groups.

Total Costs (TCI). To measure economies of scale, we chose to include the total costs of inputs used to provide the various outputs of the banks. These include wages and salaries of officers and employees, interest paid to attract funds, and the actual or estimated cost of on-premise or off-premise computer expense. We differed from other studies in two significant respects. We included interest costs because we believed them to be a legitimate input the costs of which are both significant in size and apt to differ among banks. On the other hand, their inclusion raises the question of endogeneity as an independent variable in the regression equation, resulting in possible biased estimates. Other researchers, with the exception of Murray and White, have excluded interest and concentrated on operating costs alone. On the other hand, others have included costs of supplies, materials, and related expenses to specific outputs included in their equations. We have not, believing that they are not of great importance to scale and scope economies. Because of our interest in measuring the impact of computer technology on scale and scope economies, we included estimates of computer expense in total costs.

Outputs. Four logs of outputs were included in our regression equations:

DEP — total deposits include demand deposits and time deposits, including certificates of deposit, $100,000 and over;

LOAN — loans outstanding include real estate loans made and serviced, commercial, consumer, construction, agricultural, and other loans;

INV — investments include short-term money market instruments and long-term securities held;

NBNK — non-balance sheet expense items (NBE) include safe deposit, trust, data services, and other agency expenses.

Input Prices. The logarithmic prices of labor, financial capital, and computer capital were included:

INT — interest cost of available funds was measured by dividing interest costs incurred on deposits and borrowings during the year by the average amount outstanding during the year;

WAGE — the average wages per employee (including officers but not directors) was calculated by dividing the aggregate wages paid during the year by the average number of officers and employees on the payroll during the year;

RENT – the average annual computer rental value per central processing unit (CPU) hour at prime time rates was provided by taking the bank's estimate of the current equivalent market monthly rental price for its on-premise mainframe computer multiplied by twelve divided by the reported average weekly number of CPU hours multiplied by fifty-two.

These estimates of average computer rental prices were available only for banks with on-premise mainframes. For other banks, the mean rental value per CPU hour by asset size decile for reporting banks was assigned to other individual banks groups by asset size decile so that banks not reporting estimated computer rental values were assigned an estimate that was identical to the mean rental value of the reporting banks in the corresponding size decile.

Other Variables. The average size of loans and the average deposit size were included to permit measurement of scale and scope economies in terms of the number of accounts rather than dollar volume.

AVLOAN – the dollar value of the average loans outstanding during the year divided by the average number of loan accounts during the year;

AVDEP – the dollar value of the average of deposits outstanding during the year divided by the average number of deposit accounts during the year;

OFFICE – the total number of full service, limited service offices and paying and receiving stations.

Since the focus of this study was to investigate the impact of computer technology upon scale and scope economies, only banks that utilized computers were included in the sample. Specifically, the questionnaire permitted classification of the number of responding banks into three groups in each year. See Table 2-1.

By deducting the "neither" category from the total of banks reporting answers to the computer status question, the sample of banks in our investigation was derived. Of these banks, the number of banks providing estimates of computer current market rental rates were included in a separate sample.

For more explicit information on the computation of the variables, see Appendix 2A.

Table 2-1. Sample Size.

	1982	1981	1980	1979
Computer Status				
Both on- and off-premise	240	184	174	153
On-premise only	84	79	75	103
Off-premise only	299	349	401	463
Neither	2	4	3	11
Total reporting answers	625	616	653	730
Banks not reporting answers	238	180	171	142
Total reporting banks	863	796	824	872
Bank Samples				
All bank sample	623	612	650	719
Computer bank sample	270	203	214	223

EMPIRICAL FINDINGS

Table 2-2 reports the description of data in the on-premise computer sample; Table 2-3 describes the pooled on-premise and off-premise computer sample. Deposit size ranges from $5.8 million to $2.6 billion, and total costs range from about $.5 million to $310 million per annum. One also notes the high standard deviation of the computer rental rates in both samples. The average wage rate in both samples around $18,000 per annum seems reasonable.

Table 2-4 reports the simultaneous regression coefficients in equations (2-1), (2-2), and (2-3), in each of the quartiles for 1982. Table 2-5 reports the entire sample in 1979, 1980, 1981, and 1982. In all cases the goodness of fit is unsurprisingly high. The simultaneous generalized least square regression with Shephard's cross-equation restrictions improves the fit considerably, when all these systems are compared to either linear or nonlinear ordinary least squares (OLS). One notes immediately that aggregation of the data leads to lower standard errors on the parameters, due to the larger population size in the 1982 sample. Furthermore, by inspection of Table 2-4 the coefficients appear to be quite different across quartiles. To test this proposition formally, we estimated F-statistics in each year. The

Table 2-2. Description of Data for On-Premise Computer Banks, 1982.

	Number of Banks	Mean Deposit Volume ($ millions)	Mean Annual Wage Rate	Mean Computer Rental Rate	Average Loan Size	Average Deposit Size	Mean Total Costs ($ millions)
1st Quartile	67	$ 38	$17,987	$ 30	$ 8,739.40	$3,133	$ 4
		(14.3)	(3,530)	(116)	(4,787)	(1,135)	(1.6)
2nd Quartile	68	86.3	18,046	33.8	8,518.70	3,137	9.3
		(14.3)	(2,108)	(57.1)	(3,451)	(667)	(1.7)
3rd Quartile	68	150	17,945	43.5	11,018	3,531	17
		(30.9)	(2,298)	(68)	(8,278)	(827)	(3.9)
4th Quartile	67	523	18,090	111	8,225	3,615	60
		(398)	(2,302)	(138)	(6,920)	(1,552)	(46)
1982 Sample	270	$225	$18,017	$ 54	$ 9,130	$3,354	$22.5
		(32)	(2,604)	(105)	(6,216)	(1,114)	(3.1)

Range for Sample	Minimum	Maximum
Deposits	$8 million	$2,650 million
Loans	$3 million	$1,855 million
Total Costs	$.6 million	$ 310 million

Note: Sample standard deviation in parentheses.

Table 2-3. Description of Data for Pooled On-Premise and Off-Premise Computers, 1982.

Asset Size Group (low to high)	*Number of Banks*	*Mean Deposit Volume ($ millions)*	*Mean Annual Wage Rate*	*Mean Computer Rental Rate*	*Average Loan Size*	*Average Deposit Size*	*Mean Total Costs ($ millions)*
Group #1	155	$30.1	$18,061	$26.4	$8,311	$2,979	$3.2
		(13.1)	(3,446)	(76.4)	(3,766)	(1,010)	(1.4)
Group #2	156	64.1	18,019	28.5	8,767	3,191	6.8
		(22.4)	(2,594)	(38.0)	(3,576)	(814)	(2.5)
Group #3	157	111.0	18,128	39.8	10,333	3,456	12.4
		(41.3)	(2,320)	(46.0)	(6,277)	(725)	(5.0)
Group #4	155	359.1	18,371	100.1	10,982	3,861	41.0
		(331.1)	(2,717)	(98.4)	(11,926)	(1,620)	(38.3)
All banks	623	140.8	18,145	48.6	9,600	3,372	15.8
		(210.6)	(2,795)	(75.1)	(7,280)	(1,144)	(24.3)

Range for Sample	*Minimum*	*Maximum*
Deposits	$5.8 million	$2,650.4 million
Loans	$2.8 million	$1,855.5 million
Total Costs	$ .7 million	$ 310.5 million

Note: Sample standard deviation in parentheses.

Table 2-4. Seemingly Unrelated Regression Coefficients for 1982.[a]

	Transcendental Logarithmic Equation *Dependent Variable: Log. of Total Costs*				
Independent Variables[b]	*1st Quartile*	*2nd Quartile*	*3rd Quartile*	*4th Quartile*	*Pooled Sample*
CONSTANT	.64	16.2	42.1	-1.48	8.42[c]
	(5.47)	(29.4)	(31.2)	(4.17)	(1.06)
DEP	-1.48	-8.3	-5.9	-1.09	-.36
	(2.3)	(6.6)	(3.9)	(-.61)	(.53)
LOAN	1.98	5.6	.52	1.097	.44
	(1.41)	(3.6)	(2.7)	(1.18)	(.31)
INV	1.76	2.8	2.62	2.38[c]	1.09[c]
	(1.02)	(.24)	(1.85)	(.7)	(.2)
NBNK	-.12	-.15	.68[c]	-.53[c]	-.13[c]
	(.13)	(.27)	(.28)	(.18)	(.05)
INT	2.44[c]	1.9[c]	2.7[c]	2.26[c]	2.55[c]
	(.22)	(.27)	(.3)	(.21)	(.08)
WAGE	-1.3[c]	-.88[c]	-1.58[d]	-1.24[c]	-1.4[c]
	(.22)	(.26)	(.27)	(.22)	(.07)
RENT	-.48[c]	.77	-.71	-.14	-0.05
	(.2)	(.56)	(.54)	(.17)	(0.05)
DEP^2	2[c]	4[c]	3.76[c]	1.71[c]	1.88[c]
	(.66)	(1.5)	(.73)	(.49)	(.23)
$LOAN^2$	.65[c]	1.48[c]	1.19[c]	.64[c]	.67[c]
	(.22)	(.41)	(.29)	(.19)	(.08)
INV^2	.47[c]	.81[c]	.78[c]	.39[c]	.52[c]
	(.11)	(.18)	(.14)	(.05)	(.03)
$NBNK^2$	-.001	-.003	.01	-.00	.001
	(.003)	(.004)	(.004)	(.007)	(.001)
LOAN·DEP	-1.12[c]	-2.7[c]	-1.9[c]	-.94[c]	-1.02[c]
	(.38)	(.78)	(.45)	(.29)	(.14)
INV·DEP	-.91[c]	-1.9[c]	-1.6[c]	-.8[c]	-.9[c]
	(.26)	(.53)	(.31)	(.17)	(.09)
NBNK·DEP	-.04	.05	.059	.01	-.04[c]
	(.04)	(.05)	(.04)	(.04)	(.01)

Table 2-4. continued

Independent Variables[b]	*Transcendental Logarithmic Equation* *Dependent Variable: Log. of Total Costs* 1st Quartile	2nd Quartile	3rd Quartile	4th Quartile	Pooled Sample
LOAN·INV	.44[c] (.15)	1.03[c] (.3)	.77[c] (.22)	.33[c] (.12)	.4[c] (.06)
LOAN·NBNK	.03 (.02)	.03 (.03)	-.01 (.02)	-.02 (.03)	.02[c] (.009)
INV·NBNK	.01 (.01)	.02 (.02)	.02 (.01)	.03 (.02)	.02[c] (.007)
DEP·INT	-.11[c] (.04)	.01 (.03)	-.03 (.03)	-.03 (.03)	-.06[c] (.016)
DEP·WAGE	.128 (.04)	-.005 (.04)	.03 (.03)	.02 (.03)	.055[c] (.016)
DEP·RENT	.09 (.07)	-0.08 (.09)	-0.03 (.09)	.1[d] (.05)	.082[c] (.024)
LOAN·INT	.08[c] (.02)	.02 (.02)	.002 (.02)	.02 (.02)	.037[c] (.01)
LOAN·WAGE	-.08[c] (.02)	-.02 (.02)	0.00 (0.02)	.002 (.02)	-.03[c] (.009)
LOAN·RENT	-.03 (.04)	.004 (.05)	.067 (.055)	-.06[d] (.027)	-.04[c] (.01)
INV·INT	.1[c] (.01)	.04[c] (.02)	.05[c] (.015)	.03[d] (.01)	.05[c] (.007)
INV·WAGE	-.09[c] (.02)	.05[c] (.02)	-0.05[c] (.01)	-.02[d] (.01)	-0.05[c] (.007)
INV·RENT	.04 (.03)	.04 (.04)	-0.003 (.03)	-.03 (.02)	-.04[c] (.01)
NBNK·INT	-.01[c] (.003)	.02[c] (.002)	-.009[c] (.002)	-.01[c] (.004)	-.08[c] (.001)
NBNK·WAGE	.01[c] (.003)	.02[c] (.001)	.009[c] (.002)	.009[c] (.004)	.012[c] (.001)
NBANK·RENT	-.005 (.004)	.006 (.004)	.004 (.005)	-.001 (.007)	.002 (.88)

(*Table 2-4. continued overleaf*)

Table 2-4. continued

Independent Variables[b]	1st Quartile	2nd Quartile	3rd Quartile	4th Quartile	Pooled Sample
	Transcendental Logarithmic Equation / *Dependent Variable: Log. of Total Costs*				
INT^2	.22[c]	.18[c]	.16[c]	.14[c]	.18[c]
	(.01)	(.01)	(.15)	(.016)	(.006)
$WAGE^2$	.16[c]	.16[c]	.14[c]	.1[c]	.15[c]
	(.01)	(.01)	(.01)	(.02)	(.006)
$RENT^2$	.02[c]	.007	.002	-.003	.003[d]
	(.006)	(.007)	(.002)	(.006)	(.001)
WAGE·RENT	-.003	-.008[c]	.004	.00	.00
	(.004)	(.003)	(.003)	(.004)	(.002)
WAGE·INT	.2[c]	-.188[d]	-.15[c]	-.11[c]	-.16[c]
	(.01)	(.01)	(.01)	(.02)	(.006)
INT·RENT	-.028[d]	-.004	.01	-.01[c]	-.018[d]
	(.005)	(.003)	(.003)	(.004)	(.002)
AVLOAN	-.01	.006	.001	-.018[d]	-.001
	(.01)	(.009)	(.007)	(.006)	(.003)
AVDEP	-.2[c]	-.027	-.02	.04[c]	-.02
	(.01)	(.015)	(.015)	(.01)	(.006)
OFFICE (TNO)	.002	0.000	.002	.0009[c]	-.00037[d]
	(.002)	(.015)	(.0009)	(.00026)	(.00018)
WEIGHTED R^2	.978	.93	.936	.987	.9946
N =	154	156	155	156	621

Definitions

DEP ≡ LX1; LOAN ≡ LX2; NBNK ≡ LX4; INV ≡ LX3; INT ≡ P1; WAGE ≡ P2; RENT ≡ P3.

a. This is the pooled on-premise and off-premise computer sample.
b. All variables are in logs with the exception of OFFICE.
c. Significant in two-tailed test at 1 percent level.
d. 5 percent level of significance.

Note: Standard error in parentheses. The estimated parameters in the share equations can be readily derived from the above regressions.

Table 2-5. Seemingly Unrelated Regression Coefficients, 1979-1982.

Independent Variables[a]	*Transcendental Logarithmic Equation* *Dependent Variable: Log. of Total Costs* 1979	1980	1981	1982
CONSTANT	12.9[b] (1.1)	9.048[c] (1.12)	3.75[b] (1.34)	8.42[b] (1.06)
DEP	-3.58[b] (.9)	.078 (.748)	-1.7[b] (.74)	-.36 (.53)
LOAN	2.61[b] (.55)	.21 (.43)	.91[b] (.41)	.44 (.31)
INV	1.97[b] (.33)	-.85[b] (.28)	1.8[b] (.32)	1.09[b] (.2)
NBNK	-.11[c] (.054)	-.14[b] (.05)	.01 (.07)	-.13[b] (.05)
INT	2.8[b] (.08)	2.69[b] (.079)	2.31[b] (.08)	2.55[b] (.08)
WAGE	-1.82[b] (.086)	-1.45[b] (0.086)	-.39[b] (.076)	-1.4[b] (.07)
RENT	.02 (.07)	-.108[b] (.07)	.09 (.06)	-.05 (.05)
DEP^2	2.6[b] (.51)	1.21[b] (.29)	1.62[b] (.34)	1.88[b] (.23)
$LOAN^2$	.97[b] (.18)	.43[b] (.098)	.55[b] (.11)	.67[b] (.08)
INV^2	.49[b] (.07)	.45[b] (.05)	.5[b] (.05)	.52[b] (.03)
$NBNK^2$	.004[c] (.002)	.003 (.002)	.007[b] (.002)	.001 (.001)
LOAN·DEP	-1.5[b] (.3)	-.62[b] (.17)	-.79[b] (.2)	-1.2[b] (.14)
INV·DEP	-1.01[b] (.19)	-.68[b] (.12)	-.79[b] (.14)	-.9[b] (.09)
NBNK·DEP	.06[b] (.026)	-.04[c] (.02)	-.04[c] (.025)	-.04[b] (.01)

(Table 2-5. continued overleaf)

Table 2-5. continued

Independent Variables[a]	*Transcendental Logarithmic Equation* *Dependent Variable: Log. of Total Costs* 1979	1980	1981	1982
LOAN·INV	.497[b] (.122)	.25[b] (.07)	.26[b] (.09)	.4[b] (.06)
LOAN·NBNK	.04[b] (.02)	.029[c] (.014)	.01 (.01)	.02[b] (.069)
INV·NBNK	.02[b] (.01)	.02[b] (.009)	.017 (.011)	.02[b] (.007)
DEP·INT	-.12[b] (.02)	-0.06[b] (.019)	-.06[b] (.025)	.06[b] (.016)
DEP·WAGE	.136[b] (.024)	.08[b] (.02)	.06[b] (0.018)	.055[b] (.08)
DEP·RENT	.027 (.047)	.21[b] (.03)	.049 (.028)	.082[b] (.024)
LOAN·INT	.087[b] (.015)	.05[b] (.01)	.047[b] (.015)	.037[b] (.01)
LOAN·WAGE	-.089[b] (.015)	-.04[b] (.01)	-.028[b] (.01)	-.03[b] (.009)
LOAN·RENT	-.032 (.03)	-.14 (.02)	-.039[b] (.016)	-0.04[b] (.01)
INV·INT	.069[b] (.009)	.05[b] (.007)	.059[b] (.01)	.05[b] (.007)
INV·WAGE	-.07[b] (.009)	-.05[b] (.007)	-.049[b] (.008)	-.05[b] (.007)
INV·RENT	.005 (.018)	-.067[b] (.013)	-.023 (.013)	-.04[b] (.01)
NBNK·INT	-.014[b] (.001)	-.01[b] (.001)	-.012[b] (.002)	-.01[b] (.001)
NBANK·RENT	-.001 (.003)	.002 (.002)	.006[b] (.002)	.002 (.88)
INT^2	.21[b] (.006)	.21[c] (.007)	.23[b] (.008)	.18[b] (.006)

Table 2-5. continued

Independent Variables[a]	*Transcendental Logarithmic Equation* *Dependent Variable: Log. of Total Costs* 1979	1980	1981	1982
WAGE^2	.17[b]	.145[b]	.03[b]	.15[b]
	(.006)	(.07)	(.006)	(.006)
RENT^2	.006[b]	.01[b]	.02[b]	.003[b]
	(.002)	(.002)	(.003)	(.001)
WAGE RENT	-.005[b]	.006[b]	-.005[b]	.00
	(.002)	(.002)	(.001)	(.002)
WAGE INT	-.188[b]	-.187[b]	-.15[c]	-.16[b]
	(.005)	(.005)	(.005)	(.006)
INT·RENT	-.008[b]	.02[b]	-.029[b]	-.01[b]
	(.002)	(.002)	(.002)	(.002)
AVLOAN	-.007	-.01[b]	-0.002	-.001
	(.004)	(.004)	(0.004)	(.003)
AVDEP	-0.02[b]	-.015[c]	.004	-.02
	(.007)	(.007)	(.007)	(.006)
OFFICE (TNO)	-.0003	-.0005	0.00	.00037[c]
	(.0003)	(.0003)	(.0003)	(.00018)
WEIGHTED R^2	.9943	.9931	.9933	.9946
N =	716	649	608	621

a. All variables are in logs with the exception of OFFICE.
b. 1 percent level of significance.
c. 5 percent level of significance.
Note: Standard error in parentheses.

sum of squared residuals in the restricted regression of equation (2-1) was calculated under the assumption that the coefficients were equal in each quartile. The *F*-statistics are reported in Table 2-6. In each year, we find that the hypothesis of equality in parameters across quartiles is rejected at high confidence intervals—greater than 99 percent in all the years. We thus conclude that the production function parameters vary with bank size and that aggregation of quartiles will bias measures of scale and scope parameters.

Table 2-6. Structural Stability across Bank Size.

	F Statistic	DF	Probability ($X \geqq F$)
1979	1.73218[a]	(123,540)	0.99982
1980	2.00972[a]	(123,480)	0.999999
1981	5.42579[a]	(123,436)	0.99999
1982	1.93041[a]	(123,452)	0.9999

a. Significant at 1 percent level. The critical *F* value at 1 percent for all samples is approximately 1.4.

Table 2-7. Testing for TES and Cobb-Douglas Specification.

Sample	Degrees of Freedom	CES F Statistic	Degrees of Freedom	Cobb-Douglas F Statistic
1979–1	18,144	1.475	21,144	2.198[b]
2	18,148	.8247	21,148	1.7
3	18,148	2.798[a]		
4	18,148	1.825	(21,148)	2.62[a]
All	18,684	5.896[a]		
1980–1	18,129	1.285	(21,129)	8.2[a]
2	18,131	2.3896[b]	(21,131)	3.25[a]
3	18,131	3.78[a]		
4	18,130	3.178[a]		
All	18,617	7.89[a]		
1981–1	18,118	4.606[a]		
2	18,120	3.055[a]		
3	18,121	3.837[a]		
4	18,121	4.393[a]		
All	18,576	34.1[a]		
1982–1	18,122	3.87[a]		
2	18,124	1.258	(21,124)	1.97[b]
3	18,123	4.52[a]		
4	18,124	4.30[a]		
All	18,589	13.7[a]		

a. Rejection of the null at the 1 percent level of significance.
b. Rejection of the null at the 5 percent level but not 1 percent.
Note: If the CES is rejected, then the Cobb-Douglas will also be rejected; thus we do not show Cobb-Douglas *F* statistics in these cases.

Again by inspection of Tables 2–4 and 2–5 one notes that many of the interaction terms are significantly different from zero, perhaps due to scope economies. After imposing linear homogeneity on costs with respect to prices, we tested whether the production function could be represented by homothetic technologies. The dual of the CES production function requires that each of the parameters α_{ij} and γ_{ij} be restricted to zero. The Cobb-Douglas (a special case of the CES) must also have (in addition to the above, CES restrictions) β_{ij} equal to zero (i.e., all the interaction terms are not significantly different from zero). The F-statistics for each quartile and year 1979–1982 for the pooled on-premise and off-premise computer sampler are shown in Table 2–7. In the aggregate samples the CES specification is rejected at the 1 percent level of significance in all four years. In 1981, CES is rejected at the 1 percent level in all quartiles, three out of four quartiles in 1980 and 1982, and only once in 1979. The Cobb-Douglas specification cannot be rejected at 5 percent in the second quartile of 1979 and at 1 percent in the second quartile of 1982. In 75 percent of the samples, the CES as technology is rejected, and in only one case, we could not reject a Cobb-Douglas specification at a 5 percent level of significance. Rejecting Cobb-Douglas and CES implies that cost complementarities in the production of multiple products are present. This obviously has serious ramifications for the optimal scale of bank production as well as the variety of products due to scope economies.

Economies of Scale

Table 2–8 describes ray average costs (measured at the mean of each sample) for the pooled computer sample for each quartile and aggregate each year, 1979–1982. In the aggregate samples of 1979 and 1982, one cannot reject the hypothesis that there are constant returns to scale. But in 1980 and 1981 one can significantly reject the absence of scale economies.

When the data are split into quartiles, there are significant economies of scale in the first two quartiles in all years, and during 1980 and 1981 there are significant scale economies in the fourth quartile (largest asset size). We also find no significant diseconomies of scale in any of the samples throughout the four years. These findings sharply contradict the findings of previous researchers (Gilligan,

Table 2-8. Estimates of Ray Economies of Scale.[a]

Sample[b]	*Ray Economies of Scale Estimate (RES)*	*Standard Error*	*t-Statistic (Null: RES = 1)*
1979–1	.918[c]	.021	4.26
2	.906[c]	.026	3.55
3	.986	.022	0.58
4	.987	.011	1.85
1979	.997	.005	.65
1980–1	.943[c]	.02	2.89
2	.836[c]	.034	4.79
3	1.04	.037	-1.08
4	.967[c]	.011	3.03
1980	.98[c]	.005	3.15
1981–1	.902[c]	.024	4.07
2	.759[c]	.041	17.3
3	.985	.023	0.65
4	.972[c]	.01	2.86
1981	.967[c]	.006	5.67
1982–1	.918[c]	.019	4.26
2	.906[c]	.026	3.55
3	.986	.023	0.58
4	.982	.01	1.78
1982	.996	.005	.65

a. The *t*-statistics test the null hypothesis that RES = 1.
b. Pooled off-premise and on-premise banks with linear homogeneity imposed.
c. A two-tailed test rejection of null at 1 percent level of significance.

Smirlock, and Marshall 1984; Benston, Hanweck, and Humphrey 1982) who have used aggregate data and found no economies of scale beyond banks with $50 million in deposits and moreover, have found diseconomies in large banks. The fact that the degree of scale economies depends nonmonotonically on bank size is consistent with our earlier hypothesis (see Table 2-7) that production functions differ with size.[12] In all the four years, scale economies are largest in the second quartile—0.91 in 1979, 0.84 in 1980, 0.76 in 1981, and 0.91 in 1982.

Table 2-9. Specific Scale Economies (SUR).

	On-Premise Computer Sample				On-Premise and Off-Premise Computer Sample			
	(Linear Homogeneity Imposed)							
	Deposits	Loans	Investment	Nonbanking	Deposits	Loans	Investment	Nonbanking
1982–1	.56	.25	.13	.02	.58	.23	.13	.02
2	.08	.46	.29	.04	.21	.41	.28	.03
3	.14	.46	.33	.02	.17	.47	.3	.02
4	.33	.36	.19	.06	.27	.4	.24	.04
All	.34	.40	.22	.04	.32	.39	.24	.03
All (UR)[a]	.36	.39	.21	.04	.34	.39	.23	.02
1981–1	.76	.11	-.13	.1	.59	.28	.16	.03
2	.29	.37	.23	.04	.34	.36	.22	.04
3	-.05	.56	.34	.01	.28	.36	.20	.03
4	.41	.28	.21	.03	.58	.14	.21	.02
All	.27	.4	.21	.06	.39	.33	.21	.04
All (UR)[a]	.27	.41	.2	.05	.35	.36	.21	.04
1980–1	.81	.02	.04	.05	.8	.06	.065	.023
2	.28	.41	.32	.01	.43	.39	.21	.02
3	.21	.44	.19	.005	.25	.43	.22	.03
4	.22	.46	.19	.06	.28	.44	.18	.05
All	.31	.43	.19	.04	.5	.27	.17	.04
All (UR)[a]	.33	.41	.19	.03	.056	.24	.15	.03
1979–1	.46	.26	.14	.02	.73	.16	.06	.02
2	.52	.28	.13	.05	.57	.25	.1	.03
3	.32	.41	.18	.01	.33	.4	.18	.03
4	.32	.36	.21	.05	.35	.36	.21	.04
All	.5	.29	.15	.04	.52	.29	.14	.04
All (UR)[a]	.53	.27	.14	.04	.52	.29	.14	.04

a. This sample does not impose linear homogeneity.

Table 2–9 describes the specific scale economies or the decomposition of ray average cost elasticities into activities. Fee-based banking activity appears to contribute substantially to scale economies in all the years and across all quartiles. One would suspect that banks in this size group have not expanded their nonbanking activities to the same extent as the large money market banks since the banking deregulation bills in the 1980s. In order to compete against the bigger banks, they will have to expand this activity. Overall during 1980–1982, credit expansion is the most costly for banks, although for very small banks in the first quartile, deposit expansion is far more expensive than loan expansion, during each year 1979–1982. Money market activity (investments) is the cheapest source of potential bank activity expansion in aggregate and quartile data over all the years.

Economies of Scope

Table 2–10 describes measures of scope economies in each quartile in the pooled sample. In Table 2–11 we perform *F*-statistic tests where the null hypothesis is no scope economies for any pair of the four outputs. We significantly reject the nonlinear restriction imposed by zero scope economies. This is true for all quartiles and years. Table 2–11 presents a test for the aggregate sample in 1982.[13] The above finding is consistent with Gilligan, Smirlock, and Marshall (1984) and Gilligan and Smirlock (1984) covering the years 1973–1978, who also found significant scope economies.

Examination of Table 2–10 shows that scope economies are strongest between deposits and investment as well as deposits and loans. There are also scope economies between nonbank activity and all the others, as evidenced by the negative signs in columns 4, 5, and 6. We were, however, surprised to find strong scope diseconomies between loans and investment in all years over all quartiles. There does appear to be a lessening amount of scope economies for both loan-deposits and loan-investments. The deposit-investment aggregate scope coefficient declines from 8.4 to 2.5 over the 1979–1982 period and from 11.7 to 2 for deposit-loans over the same period. One should note, however, that the pattern of scope economies is not monotonic. In fact, from 1980 to 1981 both scope coefficients increased rather than decreased. More research will have to focus on the source of scope economies.

Table 2-10. Economies of Scope, On-Premise and Off-Premise Banks, Unrestricted (UR) and Linear Homogeneity (LH) Imposed.

	Deposits Loans		*Deposits Investment*		*Loans Investment*		*Loans Nonbank*		*Nonbank Investment*		*Deposits Nonbank*	
	UR	*LH*	*UR*	*LH*	*UR*	*LH*	*UR*	*LH*	*UR*	*LH*	*UR*	*LH*
1982–1	-4.1	-18.1	-3.5	-15.5	3.9	10.6	-.21	-.06	-.20	-.07	.14	.09
2	-49.6	-59.6	-25.2	-33.8	16.8	21.4	-.8	-.8	-.41	-.45	1.2	1.3
3	-5.0	5.3	-17.1	-7.2	2.1	-.78	.36	-.93	1.8	.7	-4.1	-3.2
4	-2.1	-5.9	-3.4	-8.0	2.9	6.3	-.6	-1.2	-1.2	-1.7	.6	1.4
All	-1.2	-2.0	-1.3	-2.5	.9	1.5	-.03	-.05	-.11	-.11	0	.06
1981–1	-2.1	-7.5	-.9	-2.2	.9	1.8	.01	-.02	-.01	-.01	.03	.01
2	2.3	-36.4	-13.2	-36.8	-2.1	40.2	.67	-4.3	-2.5	-4.4	3.2	3.8
3	1.3	3.2	.8	1.5	5.3	3.4	-1.0	-.7	.67	-.37	-.5	-.8
4	-3.8	-.75	-9.6	-11.3	3.5	5.2	-.4	-.4	-1.0	-.56	1.1	.8
All	-2.3	-4.5	-3.9	-5.6	1.9	3.4	.03	.04	.04	.04	-.07	-.07
1980–1	-3.0	-17.7	-2.3	-9.8	1.3	6.1	.01	.17	.12	.10	-.16	-.28
2	-781.8	-474.1	-419.9	-301.0	314.3	262.5	-15.3	-8.5	-8.2	-5.4	20.5	9.8
3	-47.4	-6.2	-4.1	-4.3	.71	8.6	3.97	2.67	-.03	1.7	-12.2	-.86
4	-2.9	-3.4	-4.3	-4.2	1.9	2.3	-.28	-.36	-.50	-.52	.60	.62
All	-.6	-.25	-.61	-.73	.43	.40	0	0	-.10	-.09	-.06	-.03
1979–1	-25.1	-27.2	-13.7	-14.5	9.0	9.3	-.07	-.06	-.03	-.02	.09	.07
2	4.6	-6.2	1.6	-.9	1.7	3.0	-1.2	-1.8	-.41	-.3	-1.0	.6
3	-.97	-7.9	-9.8	-6.7	2.6	2.4	.35	.30	.39	.27	-1.7	-.99
4	-1.5	-1.4	-.33	-.51	.74	.80	-.14	-.16	-.28	-.25	-.22	-.12
All	-10.8	-11.7	-8.1	-8.4	5.6	5.9	-.24	-.24	-.19	-.17	.31	.31

Table 2-11. Significance of Economies of Scope, 1982, Pooled Sample.

(nonlinear least squares)

	F-Statistic		F-Statistic
1, 2 Deposit Loans	582.59[a]	2, 3 Loans Investment	344.8[a]
1, 3 Deposit Investment	582.44[a]	2, 4 Loans Nonbank	49.6[a]
1, 4 Deposit Nonbank	589.45[a]	3, 4 Investment Nonbank	59.6[a]

a. Significant at 1 percent level of significance (critical F value is 2.01). These were estimated using nonlinear least squares. The maximum number of iterations allowed was 3,000.

The above findings suggest that strong jointness in production does exist for banks, confirming the earlier findings of Gilligan, Smirlock, and Marshall (May 1984) with a 1978 sample and Gilligan and Smirlock (March 1984) in their sample covering the years 1973 to 1977.

The Substitutability of Computers in Banks

Tables 2-12 and 2-13 summarize our findings on the elasticity of substitution between all inputs as well as elasticities of derived demand. In many cases the signs are wrong; for example, the elasticity of financial capital with respect to interest costs and labor with respect to wages are of the wrong sign. There appears to be much multicollinearity in the estimated system. A rise in yields could well be reflecting a rise in asset earnings, thus producing the wrong sign. Still one of our predictions was that financial capital would be a relatively poor substitute for labor or capital due to the balance sheet restriction or limitational input effect. This seems to be borne out by the data, where the elasticity of substitution between computers and labor is far greater than between financial capital and the physical inputs.

Two elasticities seem reasonable and are generally significant throughout all quartiles and aggregates in each year. These are the own elasticity of computers and the elasticity of substitution between labor and computers. The elasticity of substitution is relatively

Table 2-12. Elasticities of Substitution (SUR) and Own Elasticities, 1982.

	On-Premise Computer Banks			*All Computer Banks*		
	Financial Capital	*Labor*	*Computer*	*Financial Capital*	*Labor*	*Computer*
Q1						
Financial Capital	.004 (.04)	.002 (.18)	-.48 (.4)	.05[a] (.02)	-.23[a] (.08)	-.27 (.22)
Labor		-.04 (.13)	1.32 (1.28)		.14[a] (.06)	1.45[b] (.76)
Computer			.09			-.09
Q2						
Financial Capital	.06[b] (.03)	-.41[b] (.21)	.65[b] (.34)	.04[b] (.02)	-.23[b] (.09)	.30 (.21)
Labor		.32 (.17)	.32 (1.4)		.18[b] (.07)	.37 (.86)
Computer			-.58[a] (.15)			-.31[a] (.07)
Q3						
Financial Capital	-.001 (.03)	.056 (.17)	-.69 (.44)	.02 (.02)	.11 (.1)	.18 (.23)
Labor		-.09 (.14)	3.54 (2.29)		.06 (.08)	2.6[b] (1.07)
Computer			-.08 (.17)			.61[a] (.07)
Q4						
Financial Capital	-.02 (.04)	.13 (.21)	-.65 (.5)	-.03 (.02)	.20 (.11)	-.28 (.27)
Labor		-.15 (.16)	3.67 (2.09)		-.20[b] (.09)	2.5[b] (1.2)
Computer			-.23 (.19)			-.25[a] (.08)
All (LH)						
Financial Capital	.03[a] (.01)	-.16[b] (.07)	-.28 (.15)	.03[a] (.008)	-.17[a] (.04)	.05 (.1)
Labor		.09[b] (.05)	2.92[a] (.59)		.1[a] (.03)	1.8[a] (.4)
Computer			-.35[a] (.07)			-.38[a] (.08)

a. Significant at the 1 percent level. b. Significant at the 5 percent level.
Note: Linear homogeneity imposed; standard (approximate) error in parentheses.

Table 2-13. Elasticities of Substitution (SUR) and Own Elasticities (LH), Pooled On-Premise and Off-Premise Computer Banks.

	Financial	*Labor*	*Computer*
		1979	
Financial	.003	-.03	.23[a]
	(.009)	(.03)	(.09)
Labor		-.001	1.05[a]
		(.026)	(.25)
Computer			-.43[a]
			(.03)
		1980	
Financial	.02	-.09[a]	-.07
	(.008)	(.03)	(.06)
Labor		.045[b]	.71[a]
		(.027)	(.21)
Computer			.11[a]
			(.03)
		1981	
Financial	-.04[a]	.17[a]	.16[a]
	(.006)	(.03)	(.04)
Labor		(-.19)[a]	1.006[a]
		(.03)	(.15)
Computer			-.31[a]
			(.032)
		1982	
Financial	.03[a]	-.17[a]	.05
	(.008)	(.04)	(.1)
Labor		.1[a]	1.8[a]
		(.03)	(.04)
Computer			-.38[a]
			(.08)

a. Significant at the 1 percent level.
b. Significant at the 5 percent level.
Note: Standard error in parentheses.

large in the aggregate sample: In 1982 it was 2.92 in the on-premise sample and 1.8 in the pooled sample with the exception of 1980. The elasticity of substitution in the aggregate samples and 1982 quartiles was not less than unity. Among larger banks the data of 1982 confirm that on-premise computer banks tend to displace labor more than off-premise computer banks. The on-premise elasticity of substitution is 2.92 (significant at the 1 percent level) as opposed to 1.8 for the pooled sample. The derived demand elasticity for computers has been relatively small and stable from 1979 to 1982. In 1982 this value was -0.35 and -0.38 in the on-premise and pooled sample respectively.

Interest Rates and Measurement Error

With the exception of Murray and White (1983), most empirical studies have excluded interest rates from the cost function. Often total costs have included interest expenses, while the interest rate is excluded from the list of independent variables. If interest rates are correlated across banks (and there is every reason to believe they are), the estimates will be biased and inefficient. If, however, one excluded interest costs in the total costs variable and interest rates as an input price, would the results we have shown here be significantly different?

Table 2-14. Scale Economies without Interest Costs.

	Q1	*Q2*	*Q3*	*Q4*	*All*
1982	.75[a]	.81[b]	.44[b]	.98	.89[a]
	(.09)	(.07)	(.16)	(.39)	(.023)
1981	.59[a]	.53[b]	.85	.81[b]	.80[b]
	(.18)	(.14)	(.15)	(.08)	(.03)
1980	1.037	.775	.42[b]	.79[b]	.89[b]
	(.103)	(.52)	(.13)	(.05)	(.03)
1979	.83	1.036	.915	.98	.91[b]
	(.11)	(.11)	(.15)	(.04)	(.024)

a. Reject Null at 2-tailed 5 percent $t = 1.98$.
b. Reject Null at 2-tailed 1 percent $t = 2.61$.
Note: LH not imposed (off-premise and on-premise pooled).
Null: No scale economies.

In Table 2-14 we reestimate scale economies without interest expenses and an interest rate. In the aggregate sample it appears that there are significant scale economies in each year. Moreover, the scale economies after 1979 were more pronounced, being as low as 0.42 in the third quartile of 1980. We simply note that ray average costs (and the regression parameters more generally) are sensitive to interest rate specifications. Earlier work ignoring interest costs—the most important component of banking costs—cannot be robust.

CONCLUSIONS

This is an embryonic stage in our attempt to understand how technology and deregulation would affect the dynamic structure of the banking system. To do this we presented a simple theory of how financial intermediation interacts with physical production. This led us to the conclusion that as long as banks choose differential liability portfolios, exclusion of interest yields from the dual cost function is a serious empirical misspecification. Moreover, in previous empirical studies interest costs are often included in total costs; thus, leaving out an interest rate in the dual specification would again bias the results, if interest costs are correlated across banks. We have shown that banking cost functions are very sensitive to an interest rate specification.

A major innovation of this study is to include a computer rental term in our econometric model. This enabled us to estimate both the derived demand elasticity for computers and the elasticity of substitution between computers and labor from 1979 to 1982. The computer price elasticity is inelastic, but the substitutability—particularly of on-premise computer banks—is very high. Any reduction in computer rental rates relative to wages would lead to significant adjustments in the banking labor force. An implication of this is that the introduction of microcomputers, a close substitute for on-premise computers, could lead to serious employment reductions.

In contrast to some other studies we estimated individual equations for each quartile between 1979 and 1982. We found that technology parameters vary significantly within bank size and across time, with no apparent pattern. Although aggregate data tended to support earlier studies showing economies of scale for small banks and diseconomies for large banks, our disaggregated panel data show a remarkably different result. We find declining ray average costs for both small and big banks, once each quartile is estimated independently. *Thus aggregation of data may understate true scale economies.*

Significant economies of scope always prevail for deposits and investments and for deposits and loans, throughout all samples. There also appears to be scope economies between nonbank activity and with deposits, investments, and loans. With recent deregulation of certain banking product lines and strong specific scale economies for "off balance sheet" bank activity, as well as scope economies, one would suspect that the system will evolve into a highly efficient supermarket system. Our findings support the hypothesis that efficiency can only be achieved by increasing banking concentration and expanding product variety. With the advent of the new telecommunication technology, interstate banking restrictions and regulations, which explicitly prevent an integration of traditional commercial banking and nonbanking activity—such as brokerage and investment banking—will restrict the realization of significant economies of scale and scope.

APPENDIX 2A
TRANSLOG DATA VARIABLES[14]

TCI—total costs of inputs: wages (35+36+37), interest (174+291+292+293+294+295+380+381+382+383), computer rental (1,079 times twelve)

SI—interest share of TCI: interest (see TCI component)/TCI

SW—wage share of TCI: wages (see TCI component)/TCI

DEP—total deposits:
146+147+148+149+150+151+262+263+264+265+266+365

LOAN—dollar value of average loans outstanding: see numerator of average loan size

AVLOAN—average loan size: dollar value of average loans outstanding (476+531+532+533+534+626+688+689+690+691) divided by the average number of loan accounts (478+536+537+538+539+631+692+693+694+695)

AVDEP—average deposit size: total deposits (see TDEPS) divided by the average number of deposit accounts (152+153+154+267+269+270+271)

INV—average annual bank investments in U.S. securities (422), tax-exempt securities and loans (423), other investments (424), federal funds sold (425), other liquidity loans (426), trading account securities (427), and purchased real estate mortgage loans not being serviced (477)

WAGE—wages per employee (including officers but not directors): aggregate salaries (35+36+37)/number of employees (31+32)

INT—interest cost of available funds (%): interest cost (174+291+292+293+294+295+380+381+382+383+)/available funds (146+147+148+149+150+151+262+263+264+265+266+363+364+365+366)

RENT—average annual computer rental value per CPU hour, prime shift only: annual computer rental value (1079 times twelve)/annual number of CPU hours (1,080 times fifty-two)

OFFICE—total number of bank offices: full service, limited service offices and paying and receiving stations (84)

NBNK—non-balance sheet expenses: safe deposit (932), nonbanking functions (agency activities, 1,038), trust department (980), data services (1,078)

NOTES TO CHAPTER 2

1. Even if no scale or scope economies exist on the supply side, scope economies on the demand side could induce greater concentration.
2. See Gilligan, Smirlock, and Marshall (1984) for a similar finding using 1978 FCA data.
3. For example, Longbrake and Haslem (1975), Koot (1978), Murray and White (1980), and more recently the replacement of the Cobb-Douglas production function with generalized (Translog) cost function models—for example, Benston, Hanweck, and Humphrey and Gilligan, Smirlock, and Marshall.
4. An exception is Murray and White (1983), who study small-scale credit unions in Canada.
5. This has immediate implications for bank deregulation outcomes. In a free supermarket system, banks could expand output in brokerage and investment banking activity without the banking deposit restriction.
6. See Christianson and Lau (1973) for a comparison between ordinary least squares estimates and SUR estimates. In the banking area, the above system has been estimated by Humphrey (1981) and Murray and White (1983).

7. Pooled cross-section/time series was ruled out at this stage since strict confidentiality by the Fed is maintained about bank participants in each sample.
8. Different levels of these variables could also affect the parameter estimates. Gilligan, Smirlock, and Marshall and Benston et al. have estimated separate equations for branch versus unit banks.
9. In comparing our estimates with Gilligan, Smirlock, and Marshall one should note that they implicitly assume stability of parameters across quartile size.
10. See, for example, Baumol (1977).
11. See Murray and White (1983) for a derivation of (2–9). In their models all variables are standardized around their mean values. In terms of estimation, equation (2–1) in our model would be identical to theirs, except for the intercept term.
12. A monotonic relationship might exist, however, on an intraquartile basis that is between banks with the same parameters.
13. An F-test is valid here since we ignore the share equations. In many of the cases the nonlinear algorithm did not converge after 3,000 iterations! In the tests of Table 2–12, convergence was achieved.
14. Numbers shown in parentheses are variable numbers assigned by the Federal Reserve Functional Cost Analysis in its Schedule Reference Listing (SRL) of items on the data tape.

REFERENCES

Baumol, William. 1977. "On the Proper Tests for Natural Monopoly in a Multi-Product Industry." *American Economic Review* 67 (December): 809–22.

Bell, Frederick W., and Neil B. Murphy. 1968. "Costs in Commercial Banking: A Quantitative Analysis of Bank Behavior and Its Relation to Bank Regulation." Research Report No. 41. Boston: Federal Reserve Bank of Boston.

Benston, George J. 1965. "Economies of Scale and Marginal Costs in Banking Operations." *National Banking Review* 2 (June): 507–49.

Benston, George J.; Gerald A. Hanweck; and David B. Humphrey. 1982. "Scale Economies in Banking: A Restructuring and Reassessment." *Journal of Money, Credit, and Banking* 14 (November): 435–56.

Benston, George J.; Allen Berger; Gerald A. Hanweck; and David B. Humphrey. 1983. "Economies of Scale and Scope in Banking." Research Paper in Banking and Financial Economies, Board of Governors of the Federal Reserve System.

Bothwell, J., and T. Cooley. 1982. "Efficiency in the Provision of Health Care: An Analysis of Health Maintenance Organizations." *Southern Economic Journal* 46 (July): 970–84.

Christianson, D. Jorgensen, and L. Lau. 1973. "Transcendental Logarithmic Production Functions." *Review of Economics and Statistics* 55 (February): 28–45.

Denny, Michael, and Cheryl Pinto. 1978. "An Aggregate Mode with Multi-Product Technologies." In *Production Economies: A Dual Approach to Theory and Applications*, edited by M. Fuss and D. McFadden, pp. 249–67. Amsterdam: North Holland.

Gilbert, R. Alton. 1984. "Bank Market Structure and Competition: A Survey." *Journal of Money, Credit, and Banking* 16 (November, Part 2): 617–44.

Gilligan, Thomas W.; Michael L. Smirlock; and William Marshall. 1984. "Scale and Scope Economies in the Multiproduct Banking Firm." *Journal of Monetary Economics* 13 (May): 393–405.

Gilligan, Thomas W., and Michael L. Smirlock. 1984. "An Empirical Study of Joint Production and Scale Economies in Commercial Banking." *Journal of Banking and Finance* 8: 67–77.

Humphrey, David B. 1981. "Economies of Scale in Federal Reserve Check Processing Operations." *Journal of Econometrics* 15: 155–73.

Koot, R.S. 1978. "On Economies of Scale in Credit Unions." *Journal of Finance* 33 (September): 1087–94.

Longbrake, W.A., and J.A. Haslem. 1975. "Productive Efficiency in Commercial Banking." *Journal of Money, Credit, and Banking* 7: 317–30.

Murray, John D., and Robert W. White. 1980. "Economies of Scale and Deposit-Taking Financial Institutions in Canada: A Study of British Columbia Credit Unions." *Journal of Money, Credit, and Banking* 12 (February): 588–70.

____. 1983. "Economies of Scale and Economies of Scope in Multiproduct Financial Institutions: A Study of British Columbia Credit Unions." *Journal of Finance* 38, no. 3 (June): 887–902.

Sealey, C.W., Jr., and J.T. Lindley. 1977. "Inputs, Outputs and the Theory of Production and Costs at Depository Financial Institutions." *Journal of Finance* 32, no. 4 (September): 1251–65.

COMMENTARY ON CHAPTERS 1 AND 2

Mark J. Flannery

In the first two chapters each set of authors has examined a different aspect of the interaction among technology, economics, and regulation in affecting U.S. bank structure and operations. In particular, Lawrence and Shay estimate bank production cost functions, on the grounds that scale and scope economies will importantly affect the ultimate extent of interstate banking. The chapter by Berger and Humphrey discusses the underlying economics of the U.S. payments mechanism. They argue that the present system incorporates large externality-induced inefficiencies, which might be overcome by the type of changes in banking structure forecast by Lawrence and Shay. My comments primarily concern Berger and Humphrey's chapter, though I try to add some perspective to Lawrence and Shay's cost estimates in my remarks.

Berger and Humphrey's chapter can be readily divided into three components. First, the authors present detailed cost estimates for alternative payment mechanisms under current technological constraints. Their numbers indicate that transactors can effect payment more cheaply (in a social sense) via electronic payments than via checks. Second, they point out that institutional factors—primarily check collection float and the way it is (not) priced—cause a market failure that leads private individuals to employ a socially excessive number of checks. If float could be reduced or priced to payors, they argue, electronic payment media would be more widely employed, and the social cost of effecting payments would fall substantially. Finally, the authors present their central thesis, which is that the spread of interstate banking will change the economics of payment medium choice in ways that substitute more efficient for less efficient payment methods.

The authors carefully document the private and social costs of alternative payment means in Table 1-4. This information, generally unavailable elsewhere, is clearly laid out and represents a real service to those of us with an interest in the area but no taste for the amount of effort required to estimate the true relative costs. Table 1-4 clearly indicates that private user costs and social resource costs differ substantially for paper-based payments. However, an important shortcoming of this table is that it ignores (by necessity) all processing costs incurred *outside* the banking industry. Table 1-4 includes the bank fees charged to make check or ACH transfers but not the payor's (or payee's) cost of sitting down, writing a check, putting a stamp on the envelope, reconciling and correcting payment errors, and so on. Implicitly, Berger and Humphrey have assumed that these other costs are identical for electronic and nonelectronic means of payment. While incorporating these costs could surely change the indicated social costs of alternative payments media in either direction, I suspect that the newness of electronic technologies makes these omitted, user-borne costs somewhat higher for ACH payments than for checks. Suppose, for example, that people are generally confident that check transfers rarely go awry, while it is believed that some percentage of electronic payments gets lost along the way. As we all know, the (time and irritation) cost of arguing with a computer can be sizable. Even a relatively small probability of such a relatively large inconvenience could therefore substantially change the perceived relative costs of paper versus electronic payments media. Unless these issues are addressed, Berger and Humphrey have not unambiguously demonstrated that electronic payments are socially more efficient.

The contention that payees fail to charge appropriately for float gives rise to a twentieth century version of Gresham's law: bad payments media drive out good. The external effect of payment medium choice occurs because payees do not discriminate, and the obvious question to ask is why no such discrimination occurs. With large transactions the means of payment is actively negotiated and reflected in the price. For example, Treasury bills generally—75 to 80 percent of the time—trade for delivery the following business day, against good funds. A purchaser may negotiate another form of payment, but the price would be adjusted accordingly. I would like to see some discussion of why payees do not negotiate similarly in smaller transactions. Perhaps people are ignorant; perhaps the record-keeping costs far outweigh the benefits; perhaps there is some sort of equilibrium among sellers—no one else discriminates, so why should I?—that is collectively rational. Once again, without some discussion of these issues, we cannot be certain that the information provided in Table 1-4 correctly reflects the total social costs of alternative payments media.

The chapter's most innovative section argues that the structural changes associated with interstate banking will change the private incentives to employ different types of payments. In particular, Berger and Humphrey claim that check truncation is more likely to occur if the number of U.S. banking firms decreases substantially, because there would be fewer parties involved in the negotiations. Increased check truncation would reduce float by about 15 percent and increase ACH processing volume with attendant reductions in unit processing costs. I was impressed with the authors' ability to model this phenomenon, and I find their conclusions provocative and thought provoking. The only caution I would offer about these conclusions is that Berger and Humphrey here assume interstate *branching*, not merely the interstate acquisition of legally separate banks by a nationally diversified holding company. Given current restrictions on the ability of holding companies to operate their subsidiary banks as a single entity, Berger and Humphrey may have overstated the extent to which this sort of interstate banking will reduce the costs of negotiating float reduction. Of course, "interstate banking" could easily include a relaxation of these restrictions, but that appears to be a somewhat separate issue.

Overall, Berger and Humphrey are to be commended for bringing a degree of economic order and serious statistical analysis to a subject area that does not easily lend itself to such things. One might quibble with some of their methods, but the authors' ability to suggest connections between bank structure and the payments system's efficiency is commendable. Their insights provide a valuable frame of reference for a book concerned with the relation between financial intermediary regulation and technological innovation.

Lawrence and Shay use functional cost analysis (FCA) data from 1979 through 1982 to investigate the technological determinants of bank cost functions. They also produce new estimates of scale and scope economies in banking, using state-of-the-art statistical methods. Anyone who has worked with these data can attest that their careful analysis reflects an impressive amount of hard work. It is also well known that bank cost functions cannot be estimated without making a number of sometimes crude or arbitrary assumptions in constructing the data set. Rather than list the latter potential causes for complaint, most of which are not unique to the present investigations, I will limit myself to two general observations.

First, I entirely agree that technology, scale economies, and scope economies are potentially crucial determinants of the nature of interstate banking in the United States. However, I remain unconvinced that the FCA data provide an appropriate sample for the purpose at hand. The banking organizations that seem most likely to determine the shape of interstate banking are large regional and national concerns. However, the largest bank in the authors' sample holds assets of $2.5 billion, and during the sample period (1979–1982), over 88 percent of FCA sample banks had total assets below $200 million. By comparison, today's major banking organizations are so much larger and more complex that one wonders if their technological characteristics could ever be inferred from the FCA sample. Moreover, the FCA data provide a rather coarse breakdown of bank services, and Lawrence and Shay have aggregated together even the available subcategories into generic "loans" and "deposits." Can we really expect to acquire reliable statistical evidence about product differentiation and innovation, which presumably entail issues such as new product development and the economics of service bundling, from such comparatively coarse data? I think this question is a troublesome one in the context of the present chapter.

My second point about Chapter 2 concerns the statistical robustness of the reported results. The authors show that their quartile sample results differ importantly from those for the overall sample. Lawrence and Shay's completeness in documenting the sensitivity of their results to disaggregation is to be applauded, but an additional dimension of disaggregation—the distinction between unit and branch banks—should be considered. Prior studies—Benston, Hanweck, and Humphrey 1982; Flannery 1983; Gilligan, Smirlock, and Marshall 1984; Mullineaux 1978)—have rejected the hypothesis that these two types of institutions have the same cost function. Are Lawrence and Shay's quartile results therefore subject to aggregation bias because they have constrained branch and unit banks to have identical coefficients within each quartile? Although I recognize the econometric and expositional difficulties associated with subdividing a sample too many times, their conclusions are novel enough to warrant further discussion and experimentation.

REFERENCES

Benston, George J.; Gerald A. Hanweck; and David B. Humphrey. 1982. "Scale Economies in Banking: A Restructuring and Reassessment." *Journal of Money, Credit, and Banking* 14 (November): 435–56.

Flannery, Mark J. 1983. "Correspondent Services and Cost Economies in Commercial Banking." *Journal of Banking and Finance* 7 (March): 83–99.

Gilligan, Thomas W.; Michael L. Smirlock; and William Marshall. 1984. "Scale and Scope Economies in the Multiproduct Banking Firm." *Journal of Monetary Economics* 13 (May): pp. 393–405.

Mullineaux, Donald J. 1978. "Economies of Scale and Organizational Efficiency in Banking: A Profit-Function Approach." *Journal of Finance* (March): 259–80.

COMMENTARY ON CHAPTERS 1 AND 2

Michael Smirlock

Chapters 1 and 2 are broadly concerned with the effects of technological innovation and regulation on the cost structure of banks. Their emphasis is different in that Lawrence and Shay (Chapter 2) address the effect of technology on overall bank costs whereas Humphrey and Berger (Chapter 1) emphasize the costs and use of electronic payment mechanisms and how their usage will respond in a deregulated banking environment. Both studies make a contribution to our understanding of the banking structure as it exists today and how this structure will respond to continued technological innovation and deregulation. In this discussion I will concentrate my comments on the chapter by Lawrence and Shay (LS).

LS examine scale and scope economies in banking with particular attention to the effect of computer technology on the cost function of commercial banks. These computer-driven innovations include electronic cash management systems and funds transfers (mostly for corporations), automated teller machines (ATMs), home banking from personal computers (such as the PRONTO system of Chemical Bank), and point-of-sales (POS) funds transfers.

The objective of LS is threefold: first, to improve upon the methodology used by other researchers in estimation of multiproduct bank cost functions. Part of this improvement inclusion of computer capital costs in the analysis, estimation of elasticities of substitution, and a redefinition of bank outputs that expands the vector of bank output. The second objective is to develop and use measures that allow determination of the effect of technological developments affecting banking upon economies of scale and scope. The third objective is to provide some insight on the impact of interstate banking.

The empirical results of the LS study seem to at times support and at times contradict previous findings, as well as providing new insights on issues not addressed in the existing literature. LS report what they consider significant scale economies even for banks up to $2.5 billion in deposits. This is in marked contrast to the multiproduct bank cost studies by Benston, Berger, Hanweck, and Humphrey (1983); Gilligan and Smirlock (1984); and Gilligan, Smirlock, and Marshall (1984), which report diseconomies of scale for bank sizes beyond $50 to $100 million in deposits. LS assert the difference is due to previous studies' aggregating their sample across all bank sizes in estimation. LS claim the cost function is not stable across bank size and base their scale economies claim on separate estimation by size quartile.

Their results support the Gilligan, Smirlock, and Marshall findings of scope economies or cost savings in the joint production of loans and deposits. It extends these findings by also reporting the existence of cost complementarities between deposits and investments, nonbank services and deposits, and nonbank services and investments. LS find diseconomies of scope between loans and investments. They also report elasticities of substitution between inputs. They find a substantial substitution between computers and labor which they view as indicative of a continued rise in the computer/labor ratio as computer costs continue to fall.

The LS chapter is to be commended for its attempt to integrate technological considerations into the examination of the structure of bank costs and for the insights it provides about bank costs in a deregulated environment. As the authors themselves state, their work is preliminary and the chapter itself is to be expanded and made more detailed. Given this, it is important to evaluate the chapter in light of previous literature and with respect to the research methodology used.

There are, of course, numerous bank cost studies. Only recently has the multiproduct nature of the banking firm been explicitly considered. Accordingly, allow me to limit the discussion of and comparison to past literature to two recent papers, both of which employ a vector of outputs rather than a single measure of output. These papers are by Gilligan, Smirlock, and Marshall (GSM) and Benston, Berger, Hanweck, and Humphrey (BBHH).

The basic set-up of the cost equation in these three papers is the same in that they all estimate bank costs using a translog specification. The advantages of this flexible form are described in the literature and need not be reviewed here. What is important and a major difference between this study and those by GSM and BBHH is in the definition of the variables that are employed in estimation: total costs, input prices, and output quantities. Whereas GSM and BBHH use noninterest expense to measure costs, LS define costs to include interest expense but to exclude "indirect" expenses such as materials, supplies, and (I believe although the authors do not say) occupancy expense. The latter are excluded based on the belief that they will not affect measurement of scale or scope economies. Although the effect on cost complementarities may be minor, the impact on measured scale economies may be substantial. This is most possible in branch banking states since branch banking is liable to appear more cost efficient than it actually is if occupancy expense is excluded.

The more important "redefinition" of costs is to include interest expense. Obviously, previous authors are aware of the importance of interest expense in the overall determination of bank costs. Omission of interest expense is both appropriate and valid when the issue being investigated concerns the bank's ability to produce output using internal resources and management. Given the emphasis on technology, this seems especially appropriate for the LS study. By including interest expense, LS seem to have changed the focus to concern the relative cost of producing output internally compared with purchasing funds. The effect on measured scale economies will depend in part on how output is measured. If dollar amounts are used, as in this study, the cost curve should flatten to the extent that there are diseconomies in internal production.

There is another difficulty in including interest expense in the measurement of total costs. These costs may be kept artificially low for some banks, due to the inclusion and reliance of deposits subject to interest rate ceilings. To the extent that these ceilings affect bank interest expense, the costs of bank expansion will be biased downward since it can be assumed the marginal funds needed for expansion will not come from liabilities covered by Regulation Q. Statements about the effect on costs of a "representative" bank doubling in size are very difficult to make in this case. Though I sympathize with the desire to control for interest expense, I am not sure that, at least in this case, it does not detract from the analysis.

The input prices used may also create difficulties in evaluating the results. The wage variable is defined as total wages divided by number of employees. Although this measure is frequently used in bank cost studies, it may be unsatisfactory because it confuses hours worked and labor quality with wage rates. A preferred measure might be some hourly wage rate in manufacturing. Bank wage rates should vary directly with wages in manufacturing markets, since within a state we would expect there to be labor mobility among occupations. Second, though LS are careful to include financial capital and computer capital, other capital costs are excluded. Such costs are also excluded from the calculation of total costs but the use of capital can change labor productivity and measurement of scale economies. Some measure of the cost of capital should be at least considered. Third, the interest cost variable is a problem beyond the fact it may be endogenous. The input price used is the *average*, not marginal, cost of funds. The latter will be higher than the former so any findings of economies of scale need to be considered in this light. A bank will not be able to double deposits without more than doubling interest expense. To the extent average and marginal are different, estimates of the cost function will be biased. Since their results suggest the inclusion of interest expense is not critical to scale economies measure, the role of interest in both dependent and independent variable measurement may not be a critical empirical issue. Fourth, the assignment of the mean computer rental price by asset size decile for banks not reporting estimated computer rental values (i.e., no on-premise system) is troublesome. The fact that these banks do not have an on-premise computer suggests that the characteristics of their production may be different. The methodology used by LS to assign computer use to these banks results in more than just a standard errors in variable problem and may create biases in the estimation.

The measurement of output may also be somewhat troublesome. LS use a vector consisting of four output measures: dollar amount of loans, dollar amount of deposits, dollar amount of investments, and something called non-balance sheet expense items (NBE) defined to include safe deposit, trust, data services, and other agency expenses. GSM and BBHH employ only loan and deposits as outputs and thus have the luxury of using the number of accounts as a measure of output. Number of accounts is a superior measure of output when interest expense is not included since noninterest costs are primarily related to the number of documents handled and customers served rather than to dollars deposited or loaned. Once interest expense is included—which is related to dollar quantities—the superiority of number of accounts over dollar quantities to measure output is not so obvious. Nonetheless, a problem exists in the authors' assertion that inclusion of an average deposit and loan account size variable permits calculation of scale and scope economies in terms of number of accounts will only be true if both dollar amounts and average account size are changed at the same time to control for changes in the size of a given account. A more important problem is the measurement of NBE. I cannot determine if this is a dollar measure, an account measure, or some other kind of measure. How this is measured is important enough to be stated clearly to make sure we are not lumping apples and oranges or expenses as outputs.

There are two final preestimation issues; the authors spend several pages arguing that financial intermediation is a major reason banks do not behave like a nonfinancial corporation. First, following Sealey and Lindley (1977), LS argue that financial capital, such as deposits, is an input in a fixed proportion production function. Thus, to expand production, a bank needs to increase financial capital. LS then asserts: "to avoid inconsistent estimation, this balance sheet constraint must be imposed on the cost structure." Regardless of whether or not one agrees with LS, it seems that to take this constraint seriously may involve more than simply putting interest expense on both sides of the cost equation (for example, some kind of simultaneous system). Second, GSM and BBHH provide evidence that branch and unit state banks are characterized by different cost functions. Two separate samples, one branch and one unit state bank, should be used. At a minimum, the authors should have tested for this.

The estimation itself is done very carefully and some of the salient results discussed earlier. Since the scale and scope economies measures are the heart of this study, they deserve more attention. An important finding is that the aggregate sample appears to suggest that there are constant returns to scale but when estimation is done by quartile size there is a finding of significant scale economies even in the largest quartile. This result is important and should be investigated further. As a test of the robustness of this finding, it would be interesting to determine if cost for the largest banks in a given quartile are approximately the same as the smallest banks in the next largest quartile. Since these banks are approximately the same size, their costs should be approximately the same. If this is not the case, then the robustness of the quartile results is questionable. Second, they include interest expense so that the LS findings may mean that, for whatever reason, bigger banks have lower borrowing costs, at least within any quartile. A special concern should be given to the largest bank size quartile, which probably includes banks ranging from $250 million to $2.5 billion. GSM and BBHH eliminate banks over $1 billion due to underrepresentation. It is worth determining if these $1 billion plus banks contain an outlier that is resulting in spurious scale economies calculations. I would suggest some technique such as those suggested by Belsley, Kuh, and Welsch (1980) to detect influential observations. Finally, no formula is given for how the extent of scope economies was measured. If the methodology is similar to that used by GSM, it is necessary to note that such estimates are imprecise and may be unstable.

These criticisms cast some doubt on the conclusions of the current LS study, and the issues discussed above should be addressed and considered prior to pursuing further research. Further, how valuable the results will be to policymakers needs to be considered. At least in the area of interstate banking, the LS findings may have little relevance since they apply mostly to smaller banks and most interstate activity involves the much larger multibillion banks.

The Humphrey and Berger (HB) chapter is more concerned with how scale economies will come into play in the use of electronic funds transfer as interstate banking increases. A major thrust of their study, which also contains an excellent discussion of the literature, is that there are significant scale economies in computer- or electronic-based check clearing and that interstate banking will increase the use of such systems. They argue, at least in part, that such a system is superior because real resources are spent by the payor to maximize float and by the payee to minimize float. But their argument ignores user costs outside the banking system, so that the advantage of electronic banking may be greater or less than that suggested by HB. A major aspect of their argument is that scale economies characterize electronic funds transfers and check clearing. This statement must be qualified. Volume of clearing is quite low, so it is not surprising scale economies are reported in existing studies. An important issue is at what point, if any, these scale economies cease and diseconomies of scale set in. Depending on the shape of the cost function, a massive change to electronic check clearing might actually increase average cost. Finally, it is questionable that interstate banking will provide private banks with incentives to promote the use of electronic banking.

Together, the Lawrence and Shay and Humphrey and Berger chapters provide a good starting point for future research on how technology and deregulation will affect the structure of the banking industry. It would be interesting if future work could combine the two chapters by explicitly examining the use of ATMs, POS payment schemes, and various payment-clearing mechanisms on the scale and scope economies of the banking industry and how the structure of interstate banking might be affected by these cost characteristics.

REFERENCES

Belsley, D.; E. Kuh; and R. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: Wiley.

Benston, George J.; Allen Berger; Gerald A. Hanweck; and David B. Humphrey. 1983. "Economies of Scale and Scope in Banking." Research Paper in Banking and Financial Economies, Board of Governors of the Federal Reserve System.

Gilligan, T., and M. Smirlock. 1984. "An Empirical Study of Joint Production in Banking." *Journal of Banking and Finance* 8 (March): 67–77.

Gilligan, T.; M. Smirlock; and W. Marshall. 1984. "Scale and Scope Economies in the Multiproduct Banking Firm." *Journal of Monetary Economics* 13 (May): 393–405.

Sealey, C., and J. Lindley. 1977. "Inputs, Outputs and the Theory of Production and Costs at a Depository Financial Institution." *Journal of Finance* 22 (September): 1251–66.

REJOINDER

Allen Berger

Responding to Mark Flannery's statement of our leaving out user-incurred costs, I would say we left out that part because there was no way in which we could measure it, but we hoped to capture it in our measure of the other costs. The second point Flannery makes is why merchants don't price discriminate. My comment on that is that consumers are irrational and may not be educated to be homoeconomicus. The point about the Canadian user checks: their per capita usage is about half that of the United States. Last, Flannery makes a valid point that externalities will not disappear by interstate banking because interstate banking will happen in the national level and not at the local level, and that is where most of the externalities lie. Michael Smirlock's point about our discovering economies of scale because of small usage is well taken, but let me mention that our numerical estimates show that a small increase in usage would lead to lower average costs.

REJOINDER

Colin Lawrence
Robert P. Shay

We are grateful to Michael Smirlock for his perceptive comments and have already begun or have planned to begin tests on most of the items he noted. There are a few matters that we cannot at present respond to or accept. These follow.

As to the omission of occupancy costs and real capital from our estimates, we have been unable to resolve the problem. As outsiders using Federal Reserve System data, we cannot identify the location of banks in our sample in order to use regional differences in rental rates (or wages).

Further, we do not agree with the criticism that our inclusion of interest expense in total costs and financial capital as an input changes the focus of our study "to concern the relative cost of producing output internally compared with purchasing funds." Our rationale for including them was to prevent the omission of a large element of total cost, namely the cost of obtaining deposits, associated with the provision of a major output of a bank. Physical capital, financial capital, and labor are all essential inputs to a bank's production function, and we believe that the rationale for dropping one would be similar to that for dropping any other. Further, there may be interaction between their prices. For example, computer-driven automated teller machines may lower the cost of obtaining deposits or affect the price paid for labor at a bank. Our whole approach is to take account of the simultaneity of all elements affecting the production function. We do not agree, or do not fully understand, how an "internal production function" can independently produce deposits without consideration of interest.

We acknowledge the existence of a simultaneity problem that makes interest expense questionable as an exogenous variable. But why remove one variable when all right-hand-side variables in the equation can be subject to the same criticism? We're certain that a simultaneous system of supply and *demand* must necessarily be estimated jointly so as to eliminate simultaneity. The problem is that as a practical matter, data on agents' demands for deposits and loans at a particular bank do not exist. Pragmatically then, one is forced to estimate the technology.

Though we agree that costing interest expense at the margin would be theoretically preferable to costing it as an average, we believe that since 1979 there have been an abundance of opportunities to arbitrage the differences between the costs of regulated deposits and regulation-free deposits (Regulation Q). This means that arbitrage would be used to bring marginal costs closer to average costs. If banks want to attract low-cost regulation-controlled deposits, they can increase marketing/advertising expenses and give away consumer goods or pay higher rates on decontrolled deposits. Arbitrage brings the costs of each liability closer to the other. So, while we are sympathetic to using a marginal cost proxy, we believe that in the presence of arbitrage and competition for funds, average yields should be a reasonable proxy for marginal costs.

As a practical matter, we dropped interest expense from total cost and the input of financial capital and found, ironically, that greater rather than lesser economies of scale occurred in most years (see Table 2-14). It thus appears that our work and all earlier work seems very sensitive to specification; hence, parameter estimates are not all robust and must be, as Smirlock insists, viewed with extreme caution when making policy proposals. Still, our work suggests that there is a high probability that scale and scope economies exist and bank structures depend on the *size* of banking operations. Aggregate estimates in all earlier work appear to be very biased, particularly the conclusion that large banks suffer from scale diseconomies.