Technical, Scale, and Allocative Efficiencies in U.S. Banking: An Empirical Investigation
Author(s): Hassan Y. Aly, Richard Grabowski, Carl Pasurka and Nanda Rangan
Source: *The Review of Economics and Statistics*, Vol. 72, No. 2 (May, 1990), pp. 211-218
Published by: The MIT Press
Stable URL: http://www.jstor.org/stable/2109710
Accessed: 23/01/2014 19:18

TECHNICAL, SCALE, AND ALLOCATIVE EFFICIENCIES IN U.S. BANKING: AN EMPIRICAL INVESTIGATION

Hassan Y. Aly, Richard Grabowski, Carl Pasurka, and Nanda Rangan*

Abstract—A nonparametric frontier approach is used to calculate the overall, technical, pure technical, allocative, and scale efficiencies for a sample of 322 independent banks. The sample was drawn from the Federal Deposit Insurance Corporation tapes on the Reports of Condition and Reports of Income (Call Reports) for the year 1986. The results indicated a low level of overall efficiency. The main source of inefficiency was technical in nature, rather than allocative. Separate efficiency frontiers were constructed to test the effect of branching. However, the distributions of efficiency measures for branching and non-branching banks were not found to be different.

I. Introduction

COST studies in banking are voluminous (see, for example, Benston (1965), Bell and Murphy (1967), Longbrake and Johnson (1975), Mullineaux (1978), Gilligan, Smirlock, and Marshall (1984), Kolari and Zardkoohi (1987), and Berger, Hanweck, and Humphrey (1987)).[1] Early studies attempted to measure economies of scale by using Cobb-Douglas production technologies. However, the assumption of lack of interdependence among outputs and restrictive functional forms that excluded U-shaped cost curves cast some doubt regarding the robustness of the results. With the advent of flexible functional forms (e.g., translog), researchers have been able to address the unresolved issues in prior research and in doing so also provide insights into cost complementarities. Yet most of these studies to date have focused on estimating cost functions and measuring economies of scale and scope with the implicit assumption that the banks being studied were operating efficiently.

Investigation of inefficiencies in banking units has received very limited attention. Sherman and Gold (1985), Rangan et al. (1988), and Ferrier and Lovell (forthcoming) are the only studies in recent years that have used a frontier approach to measure the efficiency of banks. The Sherman and Gold study used Farrell's (1957) approach to analyze the technical efficiency of branch bank operation. However, their study was limited to a very small sample and no attempt was made to decompose technical efficiency into pure technical and scale components. The Rangan et al. study is limited to measuring technical efficiency and its decomposition.

This study extends prior research by measuring allocative efficiency and determining whether there are significant differences in efficiency between unit and branch banking organization forms. In addition, this study provides insights into the effects of size, the extent of product diversity and level of urbanization on the overall, technical, and allocative measures of efficiency. The empirical results indicate that the different legal environments that prevail in unit banking and branch banking states do not result in any significant differences in the overall efficiency of banks. In general, the overall efficiency in banking is relatively low.

The nonparametric methodology used in this study is introduced in section II. Some issues pertinent to the specification of bank costs, outputs and inputs are discussed in section III and data and empirical results are discussed in section IV. A summary and conclusion follow in section V.

II. Methodology

The non-parametric programming approach used in this paper to construct measures of overall, allocative, and technical efficiency is based upon the work of Farrell (1957) as well as extensions of it (see Färe, Grosskopf, and Lovell, 1985). These measures are illustrated through the use of figures 1 and 2.

In figure 1, it is assumed that the firm uses two inputs, x_1 and x_2, to produce output y. The firm's production function (frontier) $y =$

Received for publication July 28, 1989. Revision accepted for publication September 28, 1989.

* Ohio State University, Marion; Southern Illinois University, Carbondale; Loyola University; and Southern Illinois University, Carbondale, respectively.

[1] See Clark (1988) for review of literature for Economies of Scale and Scope at Depository Financial Institutions.

FIGURE 1.—OVERALL, TECHNICAL, AND ALLOCATIVE EFFICIENCY

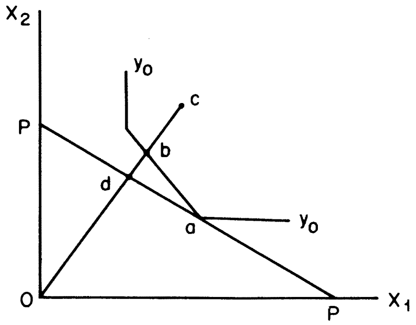
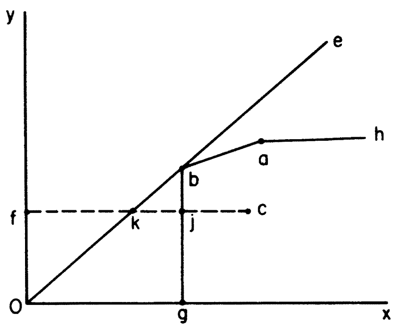


FIGURE 2.—PURE TECHNICAL AND SCALE EFFICIENCY



$f(x_1, x_2)^2$ is characterized by constant returns to scale. Let y_0y_0 and the area above and to the northeast represent all combinations of x_1 and x_2 which yield at least output level y_0. Given the technology and the input prices, represented by the slope of pp, efficient operation in production (cost minimization) occurs at point a. All cost minimizing firms are labeled as being overall efficient (OE). If point c represents a particular firm producing output y_0, then Farrell suggests measuring overall efficiency for firm c by the ratio Od/Oc. This ratio represents the potential or efficient input to actual input usage. The overall efficiency can further be decomposed into two major components: overall technical efficiency (T) and allocative or price efficiency (A). A firm is technically efficient if it is operating on the isoquant. Hence, firm c in figure 1 is technically inefficient. Farrell measures the extent of technical efficiency as $T = Ob/Oc$. Allocative inefficiencies result from choosing the wrong input combinations given input prices. For the firm at point c, A is measured by the ratio Od/Ob. The relationship between OE, T, and A is expressed as follows:

$$OE = T \cdot A. \tag{1}$$

In order to determine the overall measure of efficiency, OE, the linear programming technique is used. The first step is to calculate the minimum cost, MC, of producing the output of a particular firm. Specifically, for the multiple input, multiple output situation the following linear programming problem is solved:

$$\text{Min } px \tag{2}$$

subject to

$$y \le zY$$
$$x \ge zX$$
$$z \in R_+^k.$$

In this problem, y is the m dimensional vector of output produced by a particular firm; x is the n dimensional vector of inputs utilized by a particular firm; Y is the $(k \times m)$ matrix of outputs where k represents the number of firms; X is the $(k \times n)$ matrix of inputs; z is the vector of intensity parameters or weights attached to each of the observations or firms in the determination of minimum cost; and p is the n dimensional vector of input prices. The input values generated by the solution to the above problem represent the minimum cost level for a particular observation. Overall efficiency is then measured by

$$OE = \frac{MC}{C}, \tag{3}$$

where MC is the calculated minimum cost, and C represents the actual cost for a particular firm. This ratio corresponds to Od/Oc in figure 1.

A measure of overall technical efficiency is developed using a second linear programming (LP) problem. The LP problem is stated as

$$\text{Min } T \tag{4}$$

subject to

$$y \le zY$$
$$Tx \ge zX$$
$$z \in R_+^k.$$

In this problem T is a scalar with all of the other symbols as defined previously. In figure 1, T cor-

[2] Much of this discussion and the equations follow Førsund, Lovell, and Schmidt (1980), Färe, Grosskopf and Logan (1985), and Färe, Grosskopf, and Lovell (1985).

responds to Ob/Oc. After T and OE are calculated, allocative efficiency (A) is derived through substitution into equation (1).

Overall technical efficiency is further decomposed into measures of scale (S) and pure technical efficiency (PT). With respect to the former, if a firm is not at the optimal long-run scale of operation, i.e., constant returns to scale, the firm can hypothetically produce its current level of output with fewer inputs if constant returns to scale is attained. This is illustrated in figure 2 for the one input (x) and one output (y) case, and three observations labelled a, b, and c. A constant returns to scale production frontier is represented by Oe, which measures the optimal level of output which can be produced for given input levels. Firms either lie on or below the frontier. Thus for firm c the measure of technical efficiency is

$$fk/fc, \tag{5}$$

which corresponds to Ob/Oc in the two input case in figure 1.

In order to measure scale efficiency, the constant returns to scale assumption is dropped and a variable returns to scale frontier is developed. In figure 2 this is represented by $gbah$. The second measure of technical efficiency, pure technical efficiency (PT), is determined relative to this variable returns to scale frontier. For observation c this is written as

$$PT = fj/fc. \tag{6}$$

From the measures of technical (T) and pure technical (PT) efficiency a measure of scale efficiency is derived as

$$S = T/PT = fk/fj < 1. \tag{7}$$

This measures the proportional reduction in input usage that is achieved if the firm is operating at constant returns to scale. For constant returns to scale the value S equals 1, and all values less than 1 reflect scale inefficiency.

As illustrated in figure 2, if the measure of scale efficiency is to be calculated, then a variable returns to scale frontier must be developed and a pure technical efficiency measure (PT) derived. For the multiple output and input case this is achieved by solving the following LP problem:

$$\text{Min } PT \tag{8}$$

subject to

$$\begin{aligned} y &\le zY \\ PTx &\ge zX \\ \sum_{i=1}^{k} z_i &= 1 \\ z &\in R_+^k. \end{aligned}$$

In the above problem PT is a scalar and the requirement $\sum_{i=1}^{k} z_i = 1$[3] allows variable returns to scale. The scale efficiency measure is derived by substituting values of T and PT into equation (7). Thus, it is obvious that the scale measure is just the difference between the linear programming problems with and without constant returns to scale imposed.

If scale inefficiency exists ($S \neq 1$), the source of inefficiency is the result of operating at either increasing or decreasing returns to scale. To investigate this further, a frontier which allows for only non-increasing returns to scale, such as $Obah$ in figure 2, is constructed. This frontier is derived for the multiple output and input case by solving the linear programming problem given below.

$$\text{Min } \sigma \tag{9}$$

subject to

$$\begin{aligned} y &\le zY \\ \sigma x &\ge zX \\ \sum_{i=1}^{k} z_i &\le 1 \\ z &\in R_+^k. \end{aligned}$$

It can be proven that $\sum_{i=1}^{n} z_i \le 1$ imposes non-increasing returns to scale and σ is a scalar. If $S \neq 1$ and $\sigma = PT$ then decreasing returns to scale exists, as the σ calculation is based on a non-increasing returns to scale frontier. Alternatively, if $S \neq 1$ and $\sigma \neq PT$, then increasing returns to scale is indicated. In terms of observation c in figure 2, $S = fk/fj < 1$. Since $\sigma = fk/fc$ and $PT = fj/fc$, the result $\sigma \neq PT$ indicates increasing returns to scale.

In most prior studies utilizing a cost function approach, efficiency is attained when costs are minimized taking output or size as given. However, in this study choosing the wrong scale also constitutes inefficiency. This is due to the fact that from society's perspective, firms that operate at constant returns to scale represent the socially efficient level of operation. Thus the measures of

[3] See Afriat (1972).

inefficiency developed also incorporate inefficiency stemming from inappropriate scale.

III. Specification of Bank Costs, Outputs and Inputs

The estimation of bank costs rests on appropriate definitions and certain assumptions regarding the measurement of variables. However, the banking literature is divided on the issue of bank cost and it is not clear which variables provide good proxy measures of economic values, such as, the proxy measure of total costs. Several authors have supported the exclusion of interest expense from total costs, reasoning that interest costs are purely financial and hence are not pertinent in measuring efficiency. Others have argued that excluding interest costs disregards the process of financial technology by which deposits are transformed into loans.

Considerable disagreement also exists in prior studies on the definition of outputs and inputs for a bank. Benston, Hanweck, and Humphrey (1982) have succinctly described the problem in the following manner:

> One's view of what banks produce depends on one's interest. Economists who are concerned with economy-wide (macro) issues tend to view the banks' output as dollars of deposits or loans. Monetary economists see banks as producers of money-demand deposits. Others see banks as producing loans, with demand and time deposits being analogous to raw materials. [p. 10]

Further, the lack of a consensus in the literature on the theory of banking leaves the definition of output an unsettled issue. Hence, it is obvious that a precise definition of bank output is not possible at the present.

Prior researchers in general appear to have taken one of the two alternative approaches labeled the "intermediation approach" or the "production approach."[4] The first approach views banks as an intermediator of financial services. The outputs are measured in dollars and total costs are defined to include both interest expense and total production costs. In contrast, in the second approach the banks are viewed as producers of loan and deposit account services using capital and labor. Under the production approach, the total costs are exclusive of interest expense and the outputs are measured by the number of accounts serviced as opposed to dollar values.

This study uses the intermediation approach. Hence, interest costs are included in the total costs. The total costs are proxied by the sum of labor, capital, and loanable funds expenditures incurred by the banks in the production of outputs and services. Consistent with the intermediation approach, all outputs are measured in dollars.[5]

The inputs used in the calculation of the various efficiency measures are labor (X_1), capital (X_2), and loanable funds (X_3). Labor is measured by the number of full-time employees on the payroll at the end of the time period and capital by the book value of premises and fixed assets (including capitalized leases). Loanable funds include time and savings deposits, notes and debentures, and other borrowed funds. Capital and purchased funds are measured in thousands of dollars. P_1, the price of labor, was derived by taking total expenditures on employees divided by the total number of employees (X_1). P_2, a proxy for the price of capital, was derived by taking total expenditures on premises and fixed assets divided by book value (X_2). P_3, the price of loanable funds, was derived by taking the sum of interest expenses in time deposits and other loanable funds divided by loanable funds (X_3). The five outputs used in the study, all measured in thousands of dollars, are real estate loans (Y_1), commercial and industrial loans (Y_2), consumer loans (Y_3), all other loans (Y_4), and demand deposits (Y_5).[6]

[4] See Berger, Hanweck and Humphrey (1986) for a detailed discussion of this issue.

[5] Although it would be of interest to compare both approaches, the non-availability of number of accounts in the call report data limits the study to the application of the intermediation approach only. Thus it is presumed that the output prices faced by the banks are the same.

[6] The selection of inputs and outputs follows the study by Hancock (1986) wherein the author develops a methodology based on user costs to determine the outputs and inputs of the banking firm.

TABLE 1.—SAMPLE STATISTICS

	Mean	Standard Deviation	Minimum Value	Maximum Value
Y_1	10769	18783	50	154801
Y_2	5984	17137	0	209909
Y_3	5530	16708	52	283888
Y_4	2631	5960	0	63536
Y_5	7700	17292	0	216784
X_1	32	60	3	816
X_2	889	1766	2	16679
X_3	35680	61794	1186	830432
P_1	23.70	6.04	4.07	46.43
P_2	0.43	0.53	0.01	5.66
P_3	0.07	0.01	0.01	0.08

Note: All variables except X_1 are measured in thousands of dollars. X_1 is measured in terms of number of employees.

TABLE 2.—SAMPLE STATISTICS: EFFICIENCY MEASURE

	Mean	Standard Deviation	Maximum Value	Minimum Value
OE	.65	.18	1.00	.18
A	.87	.13	1.00	.24
T	.75	.18	1.00	.22
S	.97	.06	1.00	.70
PT	.77	.19	1.00	.22

IV. Data and Empirical Results

A. *Pooled Sample Results*

A random sample of 322 independent banks was drawn from the Federal Deposit Insurance Corporation tapes on the Reports of Condition and the Reports of Income (Call Reports) for year end 1986.[7] Some sample statistics are presented in table 1.

Summary statistics of the calculated values of the various efficiency measures that are presented in table 2 indicate a low level of overall efficiency. Specifically, the banks in this sample could have produced the same level of output using only 65% of the inputs actually used. Since the overall measure (OE) is a composite of both technical (T) and allocative (A), the relative sizes of these measures provide evidence as to the source of overall efficiency. The results indicate that the technical component is relatively more important than the allocative component as a source of overall inefficiency. Thus, inefficiency in banks may, to a greater extent, be attributed to underutilization or wasting of inputs rather than choosing the incorrect input combinations.

Looking at the components of technical efficiency, the major source of technical inefficiency is pure technical inefficiency and not scale inefficiency. The banks in this sample generally tend to lose little output due to scale efficiency. Most of the technical inefficiency exhibited by the banks stems from underutilization or wasting of inputs.

B. *Branching and Non-Branching*

In the preceding analysis, it was assumed that all banks in the sample face the same environment. As a consequence, it is implicitly assumed that all deviation from the frontier is due to inefficiency. However, the sample is composed of banks from the unit banking as well as branch banking states. These two organizational forms operate under very different legal environments and, therefore, may significantly influence the efficiency measures derived from a sample consisting of both branching and non-branching unit banks (pooled). In order to investigate this issue, the pooled sample is split into two subsamples—banks that are allowed to operate branches (212) and those that are prohibited from operating branches (110). Separate production frontiers are then calculated for each subsample. The efficiency measures discussed above are then calculated for each subsample. The results for the mean values of the efficiency measures for the separate frontiers are presented in table 3.

The issue of primary concern is whether the two sample distributions of efficiency are drawn from the same efficiency populations. The null hypothesis in this case is that the distributions of

TABLE 3.—MEAN EFFICIENCY MEASURES: SEPARATE FRONTIERS

	Non-Branching	Branching
OE	.71	.73
A	.87	.93
T	.82	.80
S	.97	.97
PT	.83	.83

[7] As the population of independent banks (independent banks are banks not affiliated with bank holding companies) is heavily skewed towards banks with a low deposit base, the random sample reflects the distribution characteristics. Further, banking holding company affiliates are excluded to isolate the effects of branching.

TABLE 4.—SUMMARY OF TESTS OF THE EFFECT OF BRANCHING ON EFFICIENCY (SEPARATE SAMPLES)

Efficiency Measure	Analysis of Variance F (Prob > F)	Median Test χ^2 (Prob > χ^2)	Wilcoxon Test χ^2 (Prob > χ^2)	Van der Waerden Test χ^2 (Prob > χ^2)	Savage Scores χ^2 (Prob > χ^2)
0	0.32	0.50	0.13	0.19	0.40
	(0.57)	(0.48)	(0.72)	(0.66)	(0.53)
A	5.35	0.50	1.67	2.24	5.83
	(0.02)	(0.48)	(.20)	(0.13)	(0.02)
T	0.34	0.03	0.64	0.77	1.11
	(0.56)	(0.87)	(0.43)	(0.38)	(0.30)
PT	0.26	0.68	0.32	0.27	0.48
	(0.61)	(0.41)	(0.57)	(0.60)	(0.49)
S	0.71	0.52	0.29	0.21	0.00
	(0.40)	(0.47)	(0.59)	(0.65)	(0.96)

efficiency measures are the same for branching and non-branching banks implying that they are drawn from the same population (environment).

The null hypothesis is tested using a variety of techniques: analysis of variance, median test, Wilcoxon Test, Van der Waerden Test, and Savage Scores Test. The analysis of variance assumes that the underlying distribution is normal and compares the samples on the basis of the within group and between group variation in efficiency. The other four tests are non-parametric in nature in that the normality assumption is not invoked. The median test compares the two sample distributions of efficiency on the basis of their central tendency, as measured by the median. The remaining three tests compare the entire structures of the distributions, not just the central tendency.[8] The results of the tests are presented in table 4.

As can be seen from table 4, for all of the efficiency measures, except allocative, the test statistics indicate that the null hypothesis cannot be rejected. As a result, it may be concluded that the differences in the distributions of the efficiency measures between the two separate samples are not significant and that they are drawn from the same population, i.e., face similar environments. Even in the case of the allocative efficiency measure only the analysis of variance and Savage Scores Test indicate rejection of the null hypothesis. The other three tests indicate that the null hypothesis cannot be rejected. In summary, the two subsamples seem to be drawn from the same efficiency population. This implies that it is appropriate to construct production frontiers by pooling (combining branching and non-branching) the data. As a consequence, the rest of the analysis in this paper deals with the efficiency results derived from the pooled sample.

TABLE 5.—MEAN EFFICIENCY MEASURES: BRANCHING VS. NON-BRANCHING BANKS (POOLED SAMPLE)

	Non-Branching	Branching
OE	.65	.64
A	.87	.85
T	.75	.75
S	.97	.97
PT	.77	.76

The analysis of the impact of branching on efficiency was discussed above within the context of separate sample efficiency results. Using the pooled sample results (summarized in table 2), the mean values for each of the efficiency measures for branching and non-branching banks are calculated. Table 5 summarizes these results.[9] The tests similar to those listed in table 4 are again applied to the efficiency results for branch-

[8] The reader is referred to Siegel (1956) for the details of the tests.

[9] Comparing the results in table 3 to those in table 5, one can see that for a number of the efficiency measures the values are higher for the separate frontier than for the pooled frontier. For example, in the pooled frontier $OE = 0.65$ while for the non-branching separate frontier $OE = 0.71$ and for the branching separate frontier $OE = 0.73$. This should not be surprising. The separate frontiers must either coincide with the pooled frontier or lie inside of the pooled frontier. This implies that the value of the efficiency measures derived from the separate frontiers must either be equal to or greater than the values derived from the pooled frontier. For a geometric proof of this, the reader is referred to Byrnes (1985).

ing and non-branching banks from the pooled sample. The results indicate that there is no statistically significant difference.[10] Thus, the efficiency results from the separate pooled samples indicate that there is no efficiency difference between branching and non-branching banks.

C. *Analysis*

Regression analysis was used to determine whether the efficiency indices derived from the pooled sample are related to the size of the bank, the product diversity of the bank, and the degree of urbanization that characterizes a bank's environment. Size is measured in two ways: the total deposits measured in thousands of dollars and the number of branches a bank operates. The measure of product diversity (D) is calculated as

$$D = -\ln \sum_{i=1}^{n} S_i^2, \tag{10}$$

where n equals the total number of different products of the firm and S_i equals the proportion of a firm's total dollar revenue or sales accounted for by the ith product. The index takes a value of zero for single product firms and increases with more product diversity.[11] The extent of urbanization is measured by using two dummy variables. The first takes on a value of one if the bank operates in an SMSA, but not in a CMSA, zero otherwise. The second dummy variable takes on a value of one if the bank operates in an SMSA that is also part of a CMSA, zero otherwise.[12]

The results of the regression analyses are presented in tables 6, 7, 8, and 9. As can be seen from table 6, the independent variables do not account for a significant share in the variation of the overall efficiency measure. However, product diversity is negatively related and CMSA positively related to overall efficiency. The first is statistically significant at the 1% and the second at the 5% level. Tables 7 and 9 show similar

[10] The results are not reported here but are available from the authors.

[11] Since the natural log of a fraction is less than zero, the resulting fraction is multiplied by minus one to convert the index into a positive number.

[12] An SMSA (Standard Metropolitan Statistical Area) is an integrated economic and social unit with a large population nucleus (fifty-thousand or more). A CMSA (Consolidated Metropolitan Statistical Area) is a large metropolitan area (over one million) which consists of more than one primary metropolitan statistical area.

TABLE 6.—REGRESSION ANALYSIS: OVERALL EFFICIENCY (OE)

Variable	Coefficient	Standard Error	t-value
Intercept	0.83	0.06	13.45[a]
Size	$1.84 \cdot 10^{-7}$	$1.38 \cdot 10^{-7}$	1.33
Product Diversity	−0.16	0.05	−3.28[a]
SMSA	−0.01	0.03	−0.19
CMSA	0.06	0.03	1.98[b]
$R^2 = .06$			
F-value = 4.91[a]			

[a] Indicates significance at the 1% level.
[b] Indicates significance at the 5% level.

TABLE 7.—REGRESSION ANALYSIS: TECHNICAL EFFICIENCY

Variable	Coefficient	Standard Error	t-value
Intercept	0.97	0.06	16.10[a]
Size	$1.92 \cdot 10^{-7}$	$1.35 \cdot 10^{-7}$	1.42
Product Diversity	−0.19	0.05	−4.02[a]
SMSA	−0.00	0.02	−0.01
CMSA	0.06	0.03	1.97[b]
$R^2 = .08$			
F-value = 6.47[a]			

[a] Indicates significance at the 1% level.
[b] Indicates significance at the 5% level.

TABLE 8.—REGRESSION ANALYSIS: ALLOCATIVE EFFICIENCY (A)

Variable	Coefficient	Standard Error	t-value
Intercept	0.85	0.04	19.70[a]
Size	$1.67 \cdot 10^{-8}$	$9.70 \cdot 10^{-8}$	0.17
Product Diversity	−0.01	0.03	0.33
SMSA	−0.00	0.02	−0.20
CMSA	0.01	0.02	0.70
$R^2 = .01$			
F-value = 0.20			

[a] Indicates significance at the 1% level.

TABLE 9.—REGRESSION ANALYSIS: PURE TECHNICAL EFFICIENCY (PT)

Variable	Coefficient	Standard Error	t-value
Intercept	0.98	0.06	16.20[a]
Size	$5.27 \cdot 10^{-7}$	$1.37 \cdot 10^{-7}$	3.85[a]
Product Diversity	−0.20	0.05	−4.20[a]
SMSA	0.00	0.02	0.11
CMSA	0.08	0.03	2.56[a]
$R^2 = .13$			
F-value = 11.39[a]			

[a] Indicates significance at the 1% level.

results with respect to technical and pure technical efficiency, except in the latter, where size is positively related and statistically significant at the 1% level. With respect to allocative efficiency, the F-statistic indicates that the multiple regression equation is not statistically significant. Analysis of the scale efficiency measure is not undertaken, given that the banks in this sample lose little output as a result of scale inefficiency.

The same regression equations are estimated using the number of bank branches as the measure of size. The results' terms of sign as well as level of significance, are very similar to those presented above.[13]

V. Summary and Conclusions

Banks in this sample were characterized by relatively low levels of overall efficiency and were, in general, more allocatively than technically efficient. Most of the technical inefficiencies observed were due to operating off the isoquant (wasting inputs) and were not scale related. On average, the banks in this sample were scale efficient. The sample included banks in branching and unit bank states. However, there was no significant difference in efficiency between the two groups. Overall and technical efficiency were negatively related to product diversity and positively related to the extent of urbanization (CMSA). In addition, pure technical efficiency was positively related to size, irrespective of whether size was measured in terms of total deposits or number of branches.

[13] The results are available from the authors upon request.

REFERENCES

Afriat, Sydney, "Efficiency Estimation of Production Functions," *International Economic Review* 13 (Oct. 1972), 568–598.

Bell, F., and N. Murphy, "Costs in Commercial Banking: A Quantitative Analysis of Bank Behavior and Its Relation to Bank Regulation," Research Report #41 (Federal Reserve Bank of Boston, 1967).

Benston, George J., "Branch Banking and Economies of Scale," *The Journal of Finance* 20 (May 1965), 312–331.

Benston, George J., George Hanweck, and David B. Humphrey, "Scale Economies in Banking: A Restructuring and Reassessment," *Journal of Money, Credit and Banking* 14 (Nov. 1982), 435–456.

Berger, Allen N., George A. Hanweck, and David B. Humphrey, "Competitive Viability in Banking: Scale, Scope and Product Mix Economies," *Board of Governors of the Federal Reserve System*, Working Paper Series (1986).

———, "Competitive Viability in Banking: Scale, Scope and Product Mix Economies," *Journal of Monetary Economics* 20 (Dec. 1987), 501–520.

Byrnes, Patricia, "Ownership and Efficiency in the Water Supply Industry: An Application of the Nonparametric Programming Approach to Efficiency Measurement," unpublished dissertation, Southern Illinois University (1985).

Clark, Jeffrey, "The Economies of Scale and Scope at Depository Financial Institutions: A Review of the Literature," *Economic Review* 73 Federal Reserve Bank of Kansas City (1988), 16–33.

Färe, Rolf, Shawna Grosskopf, and James Logan, "The Relative Performance of Publicly-Owned and Privately-Owned Electric Utilities," *Journal of Public Economics* 26 (Feb. 1985), 89–106.

Färe, Rolf, Shawna Grosskopf, and C. A. K. Lovell, *The Measurement of Efficiency of Production* (Boston: Kluwer-Nijhoff, 1985).

Farrell, M. J., "The Measurement of Productive Efficiency," *Journal of the Royal Statistical Society*, Series A, General, 120 (1957), 253–281.

Ferrier, Gary, and C. A. K. Lovell, "Measuring Cost Efficiency in Banking," *Journal of Econometrics* (forthcoming).

Førsund, F. R., C. A. Knox Lovell, and Peter Schmidt, "A Survey of Frontier Production Functions and of Their Relationship to Efficiency Measurement," *Journal of Econometrics* 13 (May 1980), 5–25.

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

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

Kolari, J., and A. Zardkoohi, *Bank Costs, Structure and Performance* (Lexington, MA: D. C. Heath, 1987).

Longbrake, W. A., and M. K. Johnson, "Economies of Scale in Banking," *The Magazine of Bank Administration* (July 1975), 32–38.

Mullineaux, Donald J., "Economies of Scale and Organizational Efficiency in Banking: A Profit Function Approach," *The Journal of Finance* 33 (Mar. 1978), 259–280.

Rangan, Nanda, Hassan Y. Aly, Carl Pasurka, and Richard Grabowski, "The Technical Efficiency of U.S. Banks," *Economics Letters* 28 (1988), 169–175.

Sherman, H. D., and Franklin Gold, "Bank Branch Operation Efficiency: Evaluation with Data Envelopment Analysis," *Journal of Banking and Finance* 9 (June 1985), 297–316.

Siegel, S., *Nonparametric Statistics for the Behavioral Sciences* (New York: McGraw Hill, 1956).