Using a Probabilistic Frontier Production Function to Measure Technical Efficiency
Author(s): C. P. Timmer
Source: *Journal of Political Economy*, Vol. 79, No. 4 (Jul. - Aug., 1971), pp. 776-794
Published by: The University of Chicago Press
Stable URL: https://www.jstor.org/stable/1830828

Using a Probabilistic Frontier Production Function to Measure Technical Efficiency

C. P. Timmer

Food Research Institute, Stanford University

This article uses linear programming techniques to "estimate" a frontier Cobb-Douglas production function for U.S. agriculture from 1960 to 1967, using the "average farm" in each state in each year as an observation. Both deterministic and probabilistic frontiers are generated and the results compared with ordinary least-squares and analysis of covariance estimates of the production function. Technical inefficiency is defined relative to the probabilistic frontier function and the extent of any inefficiency calculated for each state. Little technical inefficiency exists across states when the production function includes intermediate inputs as well as land, labor, and capital.

Introduction

The economic decision-making process can fail in two different ways. The whole core of economic theory is concerned with the first of these—the marginal revenue products of some or all factors might be unequal to their marginal costs. If this is true the allocative decision is said to be inefficient. The second source of failure is the technical production function—a failure to produce the greatest possible output from a given set of inputs means the technical decision is inefficient.

Two factors have brought the technical decision under the scrutiny of economists. First, linear programming techniques revealed a similarity between resource allocation within the firm and allocation among firms and industries. "There is both economy and additional insight to be gained by pushing the domain of study back into the firm to examine its *internal* decisions" (Walters 1963, p. 2). Second, the introduction of steady growth rather than the steady state as the desired equilibrium in economists' models raised a number of questions. How does growth start? What determines its speed? Do traditional factors of production account for it? As answers to these questions emerge, it appears that the differ-

This paper is based on my Ph.D. dissertation (Timmer 1969). A revised version has appeared under the same title (Timmer 1970).

ences between "best" practice and "average" practice in an industry are explained by the same sort of factors that explain differences in growth rates.

The extent of technical efficiency in an industry is, then, important.[1] Knowledge of the sources of any inefficiencies is doubly important. It is the purpose of this paper to (1) present a technique for measuring technical efficiency relative to a probabilistic frontier production function, (2) report the results of applying the technique to U.S. agriculture at the state aggregate level, and (3) compare the results with more traditional estimates of the production function and analysis of covariance estimates of technical efficiency. A preliminary attempt to explain the observed inefficiencies will also be reported.

If all firms faced identical factor costs and the same technical production function, and were all allocatively efficient, then relative average cost data would be sufficient to measure relative technical efficiency. Consider the simple two-factor world in figure 1. Both firms produce one unit of output, face the same relative factor costs $K'/L' = K''/L''$, and have access to the unit output isoquant production function F. If the production process is linearly homogeneous, both X_1 and X_2 are allocatively efficient. But X_2 uses more of both inputs than X_1 to produce the same level of output. Firm X_2 is technically inefficient. In this idealized world, it is easy to find a measure of this inefficiency, for $K'L'$ and $K''L''$ are isocost lines per unit of output. If X_1 is 100 percent efficient, its costs

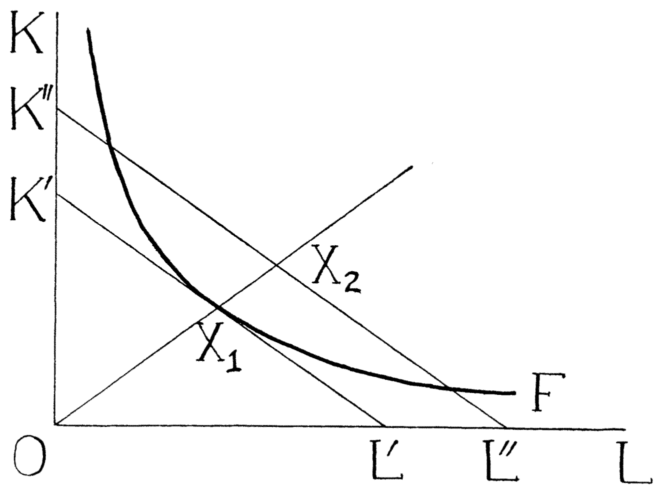

FIG. 1

[1] For further arguments and preliminary quantitative judgments on the importance of technical efficiency ("X-efficiency"), see Leibenstein 1966; Comanor and Leibenstein 1969.

serve as the basis for comparison. Then $OX_1/OX_1 = 1$ and $OX_1/OX_2 < 1$. If firm X_2 were to use precisely the same techniques as X_1, then X_2's costs could be reduced to $K'L'$. This is equal to OX_1/OX_2 $(K''L'')$. Thus the ratio OX_1/OX_2 is a measure of technical efficiency, and in this simple case it is just an index of relative costs.

Farrell (1957) generalized this idealized two-factor cost comparison to cases of n factors, differing factor costs among firms, and allocative inefficiencies as well as technical inefficiencies. The assumption of a linear homogeneous production function (LHPE) was retained. Figure 2 shows how Farrell measured each firm's technical efficiency relative to an achieved efficiency frontier. That is, AB is the envelope of the observations for all firms. No firm is able to produce a unit of output with a capital-labor input to the southwest of AB. Both firm Q and firm Q' are 100 percent technically efficient, for both produce on the unit isoquant AB. Firm P is inefficient, and the degree of inefficiency is measured by the ratio $OQ/OP < 1$. Thus all firms have a technical efficiency rating between zero and one.

Allocative efficiency can also be measured with the information in figure 2. Let DD' be the relative price line facing all firms in the industry. Then only Q' is both technically efficient and price efficient, for DD' is tangent to AB at Q'. Firm Q is technically efficient but not price efficient. The extent of its price inefficiency is measured by $OR/OQ < 1$. Since DD' is an isocost line, it is possible to produce a unit of output at a cost of only OR. Firm Q spent OQ, so the measure of its inefficiency is OR/OQ. It is clear now that Farrell's method is a step beyond simple cost comparisons. By measuring technical efficiency relative to an achieved ef-

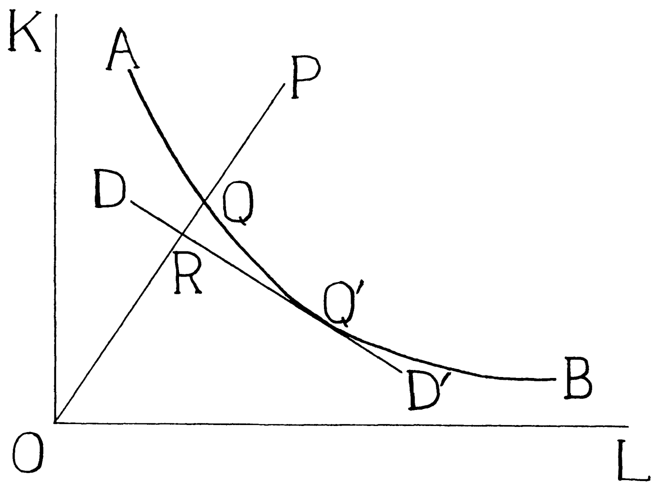

FIG. 2

ficiency frontier, Farrell was able to separate the allocative and technical decisions, something that simple cost comparisons cannot do.

Because only the extreme observations are used, the estimation of the frontier is highly subject to errors in the data. Why not use a statistically estimated *average* unit isoquant, rather than a frontier isoquant, and just measure technical efficiency relative to the average? The answer is that the frontier function, which determines "best" practice in the industry and which all firms are attempting to emulate, may *not* be a neutral transformation of the average function. The frontier production function may have entirely different factor elasticities from the average function. Comparisons of efficiencies relative to the average function will not be very useful in those cases.

The basic philosophy of the Farrell technique is used in the model to be presented here, but the details show a number of important differences. The assumption of linear homogeneity is relaxed, but at the cost of specifying a functional form, in this case Cobb-Douglas. The frontier is thus estimated in input-output space rather than input-input space. The frontier is estimated in probabilistic fashion by constraining X percent of the observations to fall outside the frontier surface. Alternative nondeterministic specifications are also investigated. The resulting model and the estimation techniques resemble those recently presented by Aigner and Chu (1968) for a somewhat different purpose.

Frontier Production Functions

Consider the usual Cobb-Douglas production function in general form,

$$y_{jt} = \sum_{i=0}^{m} x_{ijt}{}^{a_i} e_{jt} \,, \tag{1}$$

where y_{jt} = output of firm j in year t, x_{ijt} = use of factor i by firm j in year t, α_i = factor elasticities, and e_{jt} = a random error term that contains a systematic efficiency term as well. In logs (capital letters), this can be written as

$$Y_{jt} = \sum_{i=0}^{m} \alpha_i X_{ijt} + E_{jt} \,, \tag{2}$$

where one column of X_{ijt} is a vector of ones to allow for an intercept.

If all E_{jt} are constrained to one side of the estimated production surface the resulting function is an envelope—either a frontier function or an "anti-frontier" function. To be an efficient frontier, equation (2) should be estimated such that

$$\sum_{i=0}^{m} \hat{a}_i X_{ijt} = \hat{Y}_{jt} \geq Y_{jt} \,. \tag{3}$$

Only efficient firms satisfy the final equality. All others have a smaller actual output than would be achieved if they too were efficient by the standards of the frontier production function.

An infinite number of sets of $\hat{a}_i$ will satisfy equation (3). To force the estimated production surface to lie as closely as possible to the observed set of points, a minimizing constraint must be placed on some function of the sum of the resulting error terms. Although minimizing

$$\sum_{j=1}^{n} E_{jt}^2$$

would be most convenient for making comparisons with average functions, the quadratic constraint accentuates extreme observations, an undesirable feature when working with envelopes. The alternative is to minimize the linear sum of the errors, that is,

$$\sum_{j=1}^{n} E_{jt} \, .$$

Extreme observations are not unduly weighted this way.

By setting all $E_{jt} \geq 0$, equation (3) can be written as an equality:

$$\sum_{i=0}^{m} \hat{a}_i X_{ijt} - \hat{E}_{jt} = Y_{jt} \, . \tag{4}$$

The estimation technique then is to minimize

$$\sum_{j=1}^{n} \hat{E}_{jt}$$

subject to

$$\sum_{i=0}^{m} \hat{a}_i X_{ijt} \geq Y_{jt}$$

and $\alpha_i \geq 0$. In order for this to be solved by linear programming,

$$\sum_{j=1}^{n} \hat{E}_{jt}$$

must be expressed as a linear function of $\hat{a}_i$ and X_{ijt}. Then we sum equation (4) over j and solve for

$$\sum_{j=1}^{n} \hat{E}_{jt} :$$

$$\sum_{j=1}^{n} \hat{E}_{jt} = \sum_{j=1}^{n} \sum_{i=0}^{m} \hat{a}_i X_{ijt} - \sum_{j=1}^{n} Y_{jt} \, . \tag{5}$$

For any particular data set

$$\left(- \sum_{j=1}^{n} Y_{jt} \right)$$

is a constant. Any set of $\hat{\alpha}_i$ that minimizes

$$\sum_{j=1}^{n} \hat{E}_{jt}$$

for one value of $-\Sigma Y_{jt}$ will minimize for any other value, including zero, so the term can be dropped from equation (5) with no consequence. The remainder is suitable as a linear programming objective function, although for computational purposes it is desirable to divide by nt, the number of observations. Thus the arithmetic mean of the observations on the ith input, $\bar{X}_i$, is used instead of the total. In expanded form, the estimation technique requires the program to minimize $\hat{\alpha}_0 + \hat{\alpha}_1\bar{X}_1 + \ldots + \hat{\alpha}_m\bar{X}_m$ subject to

$$\begin{array}{ccc} \hat{\alpha}_0 + \hat{\alpha}_1 X_{11t} + \ldots + & \hat{\alpha}_m X_{m1t} \geq & Y_{1t} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \hat{\alpha}_0 + \hat{\alpha}_1 X_{1nt} + \ldots + & \hat{\alpha}_m X_{mnt} \geq & Y_{nt}\,, \end{array} \tag{6}$$

and $\alpha_i \geq 0$. This can be solved by any linear programming package. The vector $\mathbf{Y}_{jt}/\hat{\mathbf{Y}}_{jt}$ is the index of efficiencies, with a separate measure for each firm. These measures are averaged over time to reach a single estimate of each firm's technical efficiency, as follows:[2]

$$\text{Technical efficiency of firm } j = \frac{1}{s}\sum_{t=1}^{s}\frac{Y_{jt}}{\hat{Y}_{jt}}\,.$$

So far the discussion has been in terms of a deterministic frontier. To avoid the problem of spurious errors in the extreme observations, it is desirable to fit a probabilistic frontier. Thus equation (3) should be translated to a probability statement of the form

$$\Pr\left(\prod_{i=0}^{m} x_{ijt}\hat{\alpha}_i \geq y_{jt}\right) > P\,,$$

with P an externally specified probability—for example, 98 percent, with which the inequality is to hold (Aigner and Chu, p. 838).

The easiest way to do this is to estimate the problem in equation (6) in its entirety and determine the 100-percent efficient firms. There will be as many efficient firms as there are factors of production with $\hat{\alpha}_i > 0$,

[2] Averaging over time implies that technical efficiency is an attribute that persists over time and requires all firms to remain in the same relative position for the entire time period. This obviously does violence to the facts in a dynamic environment where, for instance, aggressive new firms gain in efficiency relative to old established firms. If the time period is long, considerable shifting about may be subsumed under the simple average. In this situation however, it is possible to break the time period into two or more subperiods and observe how the relative positions change. For further discussion of the dynamic aspects of technical efficiency, see Timmer 1970, pp. 123–34.

barring ties. This is an inevitable consequence of working in a linear world.[3] The technique is to discard the first $(100 - P)$ percent of the efficient observations until a prespecified level of P is reached. Alternatively, efficient observations might be discarded one at a time until the resulting estimated coefficients stabilize. Either way, the objections to estimating a frontier function because of data problems may be largely overcome in this fashion.

Average Production Functions

Standard statistical estimation of the "average" production function can also serve a purpose. The simple ordinary least-squares (OLS) estimate of the log linear Cobb-Douglas function serves as a basis with which to compare the frontier function. The error terms can also be used as a measure of efficiency. The analysis of covariance (AC) estimate, using separate intercept terms for each firm, can be used to examine the significance of management bias in the data. The separate firm coefficients can be used to construct an index of technical efficiency. To use AC techniques, a time series of cross section observations is needed (Mundlak 1961; Hoch 1962).

The OLS model to be estimated is

$$y_{jt} = \prod_{i=0}^{m} x_{ijt}{}^{a_i} e_{jt} . \tag{7}$$

The index of technical efficiency based on this model is constructed by averaging $\hat{e}_{jt}$ over t. This procedure assumes there is no correlation between technical efficiency and the use of factors of production. Otherwise the efficiency measure is biased due to management bias. The AC model is designed to overcome this problem by introducing separate firm intercepts in the original estimation of the production function, in the following fashion:

$$y_{jt} = a_0 a_j a_t \prod_{i=0}^{m} X_{ijt}{}^{a_i} e_{jt} , \tag{8}$$

where a_0 = overall intercept, a_j = firm intercept, and a_t = time intercept. The error term e_{jt} is presumed to be lognormally distributed with mean one, but is now free of any firm-specific or time-specific components. The vector $\mathbf{a_j}$ contains the firm effects that persist over time. These effects determine the position of each firm's production function

[3] Thus, with six factors of production there will be six efficient firms, if each factor of production has a positive output elasticity. Because the data for all eight years are pooled into a single set of observations, there is not a separate set of six efficient firms for each time period.

relative to all other firms. Thus the vector can be used as a measure of technical efficiency.[4]

The Data

All of the models discussed so far have their logical foundations in the individual firm and its entrepreneurial decision-making process. This is especially true if there is to be any attempt to explain variations in efficiency as due to managerial characteristics and ability. Aggregation beyond the firm level is likely to obscure any such impact. However, no data set containing individual firms was available in a time series of cross section observations in which farm operator characteristics were also reported.[5] The data set used here is an 8×48 matrix, where each of the forty-eight contiguous states is considered a "farm firm" and the observations are over the eight-year interval 1960–67. What will be estimated, then, is an aggregate U.S. agricultural production function, similar in philosophy and execution to those estimated by Griliches (1963*a*, 1963*b*, 1964). The results presented here will extend those of Griliches in two directions: (1) a frontier production function will be estimated, and (2) AC techniques will be used to remove management bias from the average production function.

The main source of data is *Farm Income: State Estimates, 1949–1967* (U.S. Department of Agriculture 1968). The dependent variable, Y_{jt}, is gross agricultural output in dollars of state j in year t, divided by the number of farms in state j in year t. All factors of production are also on a per farm basis. Each Y_{jt} is built up from livestock, crop, and government payments components which are deflated separately.

The labor input variable is constructed from the series in *Agricultural Statistics* for total farm employment for both family and hired workers. This series is weighted by average days worked at farm wage work for the U.S. agricultural work force in each year. No attempt is made to correct for quality differences. No annual data are available by which this can be done.

The capital input variable is a flow construct, similar to intent to that used successfully by Yotopoulos (1967). It is a current farm operating expense reported by USDA, that includes repairs and maintenance of buildings, repairs and operation of motor vehicles and other machinery, and petroleum fuel and oil used in the farm business. It is deflated by a

[4] Simultaneous-equation bias has been assumed away in all the models by stipulating that farmers make input decisions with respect to anticipated output rather than current output. This seems a reasonable assumption for most of the farm sector.

[5] Such a data set has been discovered since this article was drafted and is presently being analyzed.

crudely weighted price index for these items. This variable for repairs and operation of capital items is assumed proportional to the total capital flow variable. This assumption would be very bad if individual farm data were used because the repairs component is so lumpy. But the lumpiness should smooth out when farm data are aggregated to state data and then converted to an "average farm" basis.

Unless land area used in an agricultural production function is weighted by some measure of its productivity, the results are useless. This is especially true when technical efficiency is being measured. The efficiency ratings could easily be swamped by land quality differentials. The solution taken here is the usual one—to assume that the real estate market does a good job so that quality differences in land are reflected by differences in the (1964) sales value. Any site value captured by the sales value is assumed to be negligible. Land values are converted to a flow input of land using a 5 percent discount rate.

Actual physical quantities of each fertilizer nutrient—N, P_2O_5, and K_2O—applied in each crop year (July 1 [$t-1$] to June 30 [t]) in each state, weighted by a set of implicit price-productivity weights[6] for the 1960–67 period, are used in the production function.

The definition of the livestock variable is:

$$\begin{aligned} V_{jt} = {} & \mathrm{feed}_{jt}(W^1_{60}) + \mathrm{CLX}_{jt}(W^2_{60}) + 0.08[\mathrm{cattle}_{jt}(P^1_{60}) + \mathrm{hogs}_{jt}(P^2_{60}) \\ & + \mathrm{sheep}_{jt}(P^3_{60}) + \mathrm{chickens}_{jt}(P^4_{60}) + \mathrm{turkeys}_{jt}(P^5_{60})] \,, \end{aligned} \tag{9}$$

where feed_{jt} = current feed expenses for state j in year t, CLX_{jt} = current livestock expenses (purchases) for j, t, W^i_{60} = price deflators for feed and livestock expenses, cattle_{jt}, . . . turkeys jt = numbers of such animals on farms January 1 of year t in state j, and P^i_{60} = average value per animal in 1960 for the whole country.

The 8 percent discount rate to convert livestock capital values into flows is the same as that used by Griliches (1964, p. 967). The two different rates for livestock (8 percent) and for land (5 percent) presume that farmers are not equating marginal returns from the last dollar invested in land and livestock. But some account must be taken of the trade-off between average return and variance. Since the variance seems to be higher for returns to livestock than for land, farmers must earn a higher average rate of return from livestock to be indifferent between investing in livestock and investing in land.

The seed and miscellaneous variable is the sum of two current expense items. The seed component also includes minor amounts for bulbs, plants, and trees. It is deflated by the seed component of the prices paid by

[6] The weights are N = 1.779, P_2O_5 = 1.355 and K_2O = 0.984. They were derived from the following regression equation: total price per ton $= A + P_{\mathrm{N}}(\%\mathrm{N}) + P_{\mathrm{P_2O_5}}(\%\mathrm{P_2O_5}) + P_{\mathrm{K_2O}}(\%\mathrm{K_2O})$. The data were prices paid by farmers for different analysis fertilizer—for example, 10-10-10, 10-15-10, etc.

farmers index. The miscellaneous component includes a large variety of items and is deflated by the index for all commodities bought. Of special relevance in the miscellaneous index are charges for pesticides, electricity, irrigation, and veterinary services and medicines. These are "modern" inputs that may account for a substantial amount of productivity differentials.

The Empirical Production Functions

Given the nature of the data, it is perhaps surprising that there are any empirical production functions at all. But table 1 shows that there are.

Equation (I) of table 1 reports the simplest possible average production function estimated by ordinary least squares with all six factors of production. The R^2 of 0.97 shows that the fit is very good. These six factors account for all but 3 percent of the variation in aggregate U.S. agricultural output across states and time. All of the coefficients are highly significant. None of the coefficients are very surprising, either a priori or in light of Griliches's results. The ratios of marginal revenue product to marginal cost are about 1 for labor (1.17) and livestock (1.05), much less than 1 for land (0.29), somewhat greater than 1 for seed and miscellaneous (1.62), and much greater than 1 for capital (3.76) and fertilizer (4.86). The implications are that in the 1960–67 period farmers used too much land and not enough capital and fertilizer, although the magnitude of disequilibrium implied is suspicious.

Equation (II) of table 1 attempts to remove management bias by in-

TABLE 1

PRODUCTION FUNCTION COEFFICIENTS

(t-VALUES)

	EQUATION				
FACTORS	(I) OLS	(II) AC	(IIIa) LP_{100}	(IIIb) LP_{98}	(IIIc) LP_{97}
Constant	1.7350 (53.8)	$a_t a_j$	1.6693	1.8578	1.8828
Labor	0.1919 (6.7)	0.1231 (2.7)	0.6015	0.3287	0.2679
Capital	0.3726 (11.7)	...	0.4887	0.3689	0.4842
Land	0.0458 (4.2)	0.3443 (4.2)	...	0.0298	0.0099
Fertilizer	0.1484 (16.0)	0.0481 (2.2)	0.1334	0.1428	0.1693
Livestock	0.2510 (19.5)	0.3103 (8.3)	0.2347	0.2045	0.1885
Seed and miscellaneous	0.1579 (5.4)	0.1222 (2.8)	0.1043	0.2243	0.1712
R^2	0.970	0.994	...	...	...

troducing separate firm (and time) intercepts. There are three very striking changes in the elasticities of output:

1. The output elasticity of capital becomes completely insignificant and capital is dropped from the production function altogether. The direction of change is certainly as expected, although the magnitude is surprising. The zero elasticity is presumably more a function of the data quality than of the true productivity of capital.

2. The output elasticity of fertilizer drops very sharply as well, but fertilizer remains in the production function at a significant level. The results confirm the supposition that "good" management and heavy fertilizer applications go together. It should be noted that since the elasticity of output for fertilizer in the AC function is only about one-third that in the OLS function (equation [1]), the marginal revenue product is reduced correspondingly. The ratio of marginal revenue product to marginal cost for fertilizer was almost 5 in equation (I), a value almost identical to one reported by Griliches for a similar function but different data (1964, pp. 968–69). Elimination of management bias reduces the ratio to 1.57. This value is still high enough to suggest some disequilibrium in the use of fertilizer, and thus to explain its rapid growth in consumption, but it is not so high as to strain one's belief in the rationality of American farmers.

3. The coefficient of land in the AC production function becomes very large relative to other inputs and relative to its previous size in the OLS function. The implied disequilibrium shifts from too much land being used to not enough. The ratio of marginal revenue product to marginal cost is greater than two. This may be caused by constraints imposed by government restrictions on acreage.[7]

The last change of any significance is a drop in the labor coefficient. The size of the decline is not large, but it is sufficient to reduce the ratio of marginal revenue product to marginal cost from 1.17 with the OLS function to 0.75 in the AC function. The implications are that farmers are using too much labor at the going wage rate, and the exodus of farm labor to the cities confirms that a disequilibrium exists.

[7] The implication of the high coefficient for land in the AC production function is that "better" farmers use relatively less land than "worse" farmers. This may be accounted for by some systematic error in the land input data, for example, small farms appearing smaller than they are in fact because site value is neglected. Alternatively, smaller farms may be "better" because of more intensive cropping patterns. The direction and extent of disequilibrium implied by the AC land coefficient is in general agreement with some theoretical results generated by Hoch (1962) when acreage restrictions were imposed on his model.

All of the differences reported between OLS and AC functions should perhaps be viewed as qualitative rather than as concretely quantitative. Analysis of covariance may have some tendency to bias the estimated elasticities of output downward if there are errors in the variables.[8] The separate intercepts remove the firm specific means as information, leaving only the variance about the mean to estimate the coefficients. Any errors are thus magnified. A large sample would seem to dampen this effect.

Equations (IIIa–IIIc) report the results of fitting the log linear Cobb-Douglas production function using the linear programming model outlined above. These are frontier production functions. Equation (IIIa) labeled LP_{100}, is the result of fitting a deterministic frontier function to the same data as used in equation (I). The results are very strange, especially the very large labor coefficient. Before reading much significance into these results, it is wise to see if the extreme observations are so subject to error that the results are meaningless. Thus equations (IIIb) and (IIIc) (LP_{98} and LP_{97}) report what happens as the first 2 percent of the "efficient" firms (seven observations) from equation (IIIa), and then the next 1 percent (four more observations) are removed from the data deck.

A rather remarkable transformation takes place. With just 2 percent of the observations removed, the frontier equation looks remarkably like the OLS average function. The similarity remains when another 1 percent of the data are discarded—that is, the coefficients seem to have stabilized. The coefficients are, with minor exception, very similar to those of the analogous average function. The most obvious difference is that the intercept of the frontier function is about 14 percent higher (when the antilogs are compared) than for the average function. This is as it should be. Except for the labor coefficient, the factor elasticities are so nearly identical as to be considered the same. There are no tests of significance for this statement, however.

The 40–70 percent rise in the labor coefficient is somewhat puzzling. Apparently, the frontier farms use relatively less labor than the average farm; consequently, the marginal productivity of labor is higher. This may be related to the high capital intensity associated with "good" management that was discovered in the AC results. The frontier firms seem to be technically efficient at the expense of allocative efficiency: they use too little labor and too much capital. This may be an example of "technological man" appearing in the guise of the modern farmer (Brewster and Parsons 1946).[9]

In summary then, the frontier production function seems to have

[8] This problem was pointed out to me by Zvi Griliches. No published reference seems to be available.

[9] The referee suggests that the frontier firms may also employ higher quality labor. The effect of this would not be captured by the crude labor variable used in the analysis.

shifted almost neutrally outward from the average function, with the possible exception that the labor elasticity of output may have increased by about one-half.

Technical Efficiency

The production function estimates are interesting in their own right, but they are only a means to an end. Technical efficiency is the Holy Grail in this quest. The reason for estimating the production functions at all is to find the "right" way to correct for differential use of the factors of production. Otherwise, there is no way to judge one state's performance relative to that of another when different factor amounts and proportions are used.

Three different vectors of technical efficiency are generated from the production function results of table 1; they are reported in table 2, with rankings for each vector. The OLS residuals efficiency vector uses equation (I) as a basis for calculations. The process is in two steps: first, the production function is estimated, with the results as shown in table 1; second, the residuals from the first stage are regressed on a set of forty-eight state dummy variables. The ratings shown in table 2 are the coefficients of these dummy variables. The regression on the state dummy variables is highly significant—75 percent of the variation in the residuals is explained. But this vector of efficiencies is subject to management bias. The collinearity of management with several of the "real" factors, specifically not accounted for in the two-step procedure just outlined, allows the estimated elasticity of these factors to capture some of management's contribution. These "average" efficiency ratings are thus biased upward to some extent.

The AC efficiency vector is simply the firm intercepts of equation (II). These ratings are not the antilogs—the rankings are more interesting than the actual values. For comparison, however, antilogs have been taken of the intercepts for the most "efficient" state (Maine, 9.24) and the least "efficient" state (Oklahoma, 3.92). The Maine intercept is 2.36 times higher. The suspicion that analysis of covariance might bias the production function coefficients downward seems justified, for it is unreasonable that such large neutral shifts occur in the production functions of different states. The separate state intercepts have captured a substantial proportion of the impact of differential use of inputs, leaving very little for the factor elasticities to explain. The AC estimates of technical efficiency are thus biased downward.

The most satisfactory efficiency estimates are those calculated from equation (IIIb), the 98 percent frontier function. The high degree of technical efficiency at the state level and relative to six factors of produc-

TABLE 2

TECHNICAL EFFICIENCY RATINGS

State	OLS Residuals		AC Firm Intercepts		98% LP	
	Efficiency Rating	Rank	Efficiency Rating	Rank	Efficiency Rating	Rank
South Dakota....	0.1423	4	1.736	17	0.991	1
Iowa............	0.0936	7	1.657	25	0.986	2
North Dakota....	0.1560	2	1.967	4	0.984	3
Florida..........	0.2117	1	1.825	14	0.978	4
North Carolina...	0.1536	3	1.967	3	0.976	5
Delaware........	0.0752	9	1.861	11	0.970	6
Montana........	0.0826	8	1.651	26	0.965	7
Illinois..........	0.0362	15	1.658	24	0.963	8
Colorado........	0.0263	18	1.607	35	0.960	9
New Mexico.....	0.0545	13	1.453	44	0.956	10
Alabama........	0.0446	14	1.694	22	0.955	11
Kentucky........	0.1000	6	1.701	20	0.951	12
Connecticut.....	0.0654	11	1.902	5	0.948	13
California.......	0.0707	10	1.613	33	0.945	14
Nebraska........	−0.0270	30	1.633	27	0.945	14
Maine..........	0.1110	5	2.223	1	0.945	14
Kansas..........	0.0168	22	1.605	36	0.941	17
Wyoming.......	−0.0048	25	1.491	43	0.939	18
Georgia.........	0.0030	24	1.812	15	0.939	18
Vermont........	0.0216	21	1.834	13	0.934	20
Mississippi......	0.0581	12	1.688	23	0.932	21
Arkansas........	0.0264	17	1.631	28	0.928	22
New Hampshire..	0.0244	20	1.893	7	0.928	22
Massachusetts...	0.0340	16	1.976	2	0.923	24
Minnesota.......	−0.0314	33	1.760	16	0.922	25
Texas...........	−0.0296	32	1.444	45	0.921	26
New Jersey......	−0.0128	26	1.870	10	0.920	27
Wisconsin.......	0.0168	22	1.876	9	0.920	27
Oklahoma.......	−0.0403	35	1.365	48	0.920	27
Missouri........	−0.0326	34	1.579	37	0.917	30
Indiana.........	−0.0681	38	1.618	31	0.916	31
Idaho...........	−0.0237	28	1.702	19	0.913	32
New York.......	−0.0258	29	1.845	12	0.909	33
Arizona.........	−0.0956	39	1.625	29	0.906	34
Washington.....	−0.0472	36	1.616	32	0.903	35
Nevada.........	−0.1182	43	1.360	47	0.902	36
South Carolina...	0.0256	19	1.895	6	0.898	37
Oregon..........	−0.0568	37	1.548	40	0.896	38
Louisiana.......	−0.0282	31	1.576	38	0.890	39
Utah...........	−0.1136	41	1.432	46	0.889	40
Rhode Island....	−0.1149	42	1.891	8	0.887	41
Pennsylvania....	−0.1129	40	1.700	21	0.884	42
Maryland.......	−0.1632	48	1.522	42	0.883	43
Ohio............	−0.1250	47	1.570	39	0.880	44
Tennessee.......	−0.0181	27	1.608	34	0.880	44
Michigan........	−0.1183	44	1.707	18	0.854	46
Virginia.........	−0.1232	46	1.625	30	0.848	47
West Virginia....	−0.1187	45	1.531	41	0.810	48

tion is readily apparent from table 2. Three quarters of the states have measured efficiencies within 10 percent of the frontier (that is, ratings above 0.900). The least efficient state (West Virginia) is less than 20 percent away. And if differences in cropping patterns, crop-livestock ratios, and more accurate soil-climate productivity differentials could be introduced, the measured inefficiencies would be even smaller.

There is a nagging problem with this index of efficiencies. If the frontier function is almost a neutral transformation of the average function, should not the two efficiency vectors be similar? The answer is yes, and the simple correlation between the two is 0.89, as is seen in table 3. The other correlations are quite low. The problem is this: the OLS function suffers from management bias. Does the similarity of the frontier function to the OLS function mean the frontier suffers from management bias? What does management bias mean in the frontier context? These are difficult and as yet unanswered questions.

Explaining Technical Efficiency

What has happened? It is important to remember that the low degree of technical inefficiency among states is relative to a six-factor production function. These six factors include fertilizer, capital, and a seed and miscellaneous variable that contains many inputs. "Good" farmers use these inputs in large quantities and achieve large outputs. But the production function itself makes this distinction.

What has happened then is that differences in static technical efficiency largely disappear when measured at the state level. The result should not be too surprising. The residual of dynamic technical change also disappears under similar circumstances—when inputs are measured "properly" and all relevant factors, including intermediate ones, are included in the production function (Griliches 1963*b*).[10]

TABLE 3

Correlation Matrix of Efficiency Indexes

	OLS Residual	AC Firm Intercepts	98% LP
OLS residual	1.00	...	...
AC firm intercepts	0.45	1.00	...
98% LP	0.89	0.31	1.00

[10] Griliches found that the residual of technical change over time resulting from a simple capital and labor production function could be "explained" in terms of improved quality of labor and increased use of intermediate inputs. Similarly, differences in simple efficiency measures such as output per acre or output per man-year across states can be "explained" by the same sort of factors.

Although there is little inefficiency to explain, about half of it seems to be due to definitional and measurement problems in the variables. Equation (10) reports the most successful attempt to explain the efficiencies calculated from equation (IIIb) (technical efficiency determined from the linear programming model with 98 percent of the observations included [T.E. LP_{98}]) in terms of a variety of potential factors.

$$\begin{aligned} \text{T.E.}_{\text{LP}_{98}} = \underset{(9.2)}{0.8377} + \underset{(1.3)}{0.4386}\, X9 + \underset{(2.4)}{0.1329}\, X19 - \underset{(-4.1)}{0.6764}\, X20 \\ - \underset{(-0.9)}{0.0468}\, X21 + \underset{(1.7)}{83.2992}\, X25 . \quad R^2 = 0.48 \end{aligned} \tag{10}$$

The following are significance levels for various t-values: 0.68, 50 percent, 1.68, 90 percent, 2.02, 95 percent, 2.42, 98 percent, 2.70, 99 percent.

The most significant variable is $X20$, the number of days worked off the farm by the farm operator in 1964 (the year of the last Agricultural Census). More days worked off the farm means lower efficiency. Variable $X20$ is probably correcting for some bias introduced by the labor variable in the original production function. The labor input was a crude measure of man-days worked in agriculture, but correction for time worked by the number of actual laborers present was only on an annual basis for the entire country. There was no state-specific weighting for proportion of time spent in farm work. Variable $X20$ makes this correction at the efficiency stage rather than at the more appropriate level, that is, where the input variables were constructed.

Next in significance is the relative number of tenants of all types who were farm operators ($X19$) in 1964. The share-cropping tenure pattern in the South does not dominate tenant farming for the entire United States. In fact six Corn Belt states (Indiana, Illinois, Iowa, South Dakota, Nebraska, and Kansas) had a higher proportion of tenant farmers (0.25) in 1964 than nine Southern states (Virginia, North Carolina, Tennessee, South Carolina, Georgia, Alabama, Mississippi, Arkansas, and Louisiana) where the proportion was 0.21. The reason $X19$ is so significant might then be associated with the extra effort and motivation of the young tenant farmers in the non-southern states who are attempting to save enough to buy their own farms. This extra effort is not captured by the labor variable.

Variable $X25$, 1964 Soil Bank payments, is significant at only the 90 percent confidence level. The output variable used in the production function included a component for government payments. In general, these payments are on an output basis, that is, the more cotton produced the higher the payments. But some payments, primarily those under the Soil Bank Program, are made for not producing. Those states where the Soil Bank payments form a significant proportion of measured output (see table 4) should then show a high degree of technical efficiency. They are able to "produce" output without inputs.

Four of the seven states whose payments were more than 1 percent of gross output are ranked in the top ten in efficiency. If nothing else, part of the mystery of North and South Dakota's high efficiency ratings seems to be resolved.

Variable $X9$, the proportion of farm operators 45–54 years old (in 1964), probably serves as a proxy for something else. The likelihood that a farmer of this age may have one or more mature sons working the farm with him and sharing the managerial duties is suggestive.

The last variable that yields even a hint of meaning is $X21$, the proportion of farm operators who are Negro. The sign is negative, that is, the higher the proportion of Negro farm operators, the lower the efficiency. Variable $X21$ is also correlated with a number of other factors, as is shown in table 5.

The population of Negro farm operators is not evenly distributed with respect to age and education. There tend to be few young Negro farm operators and many over age fifty-five, when productivity seems to diminish rapidly. They tend to have less than seven years of schooling,

TABLE 4

Soil Bank Payments and Technical Efficiency

State	Soil Bank Payments (% of Gross Output)	98% LP Efficiency Rank
North Dakota	2.8	3
New Mexico	2.4	10
South Carolina	1.8	37
South Dakota	1.7	1
Oklahoma	1.5	27
Colorado	1.1	9
Georgia	1.1	18

TABLE 5

Correlation of Proportion of Negro Farm Operators with Other Variables

Variable Name	Simple Correlation
$X7$, age class 25–34	−0.37
$X8$, age class 35–44	−0.49
$X10$, age class 55–64	0.50
$X11$, age class 65 and over	0.31
$X12$, 0–4 years elementary education	0.75
$X13$, 5–7 years elementary education	0.69
$X16$, 4 years high school	−0.61
$X17$, 1–3 years college	−0.34
$X18$, 4 or more years college	−0.18
$X19$, number of tenants	0.47

with a decided lack of completed high school and college experience. And Negro farm operators tend to be tenants (primarily share-croppers). It seems reasonable to suggest that the negative efficiency coefficient for Negro farm operators (of questionable significance at that) is a matter of opportunity and not of ability.

The Welfare Impact

Judging from the small degree of technical inefficiency observed at the state level, the welfare losses from this particular failure of the decision-making process seem to be quite small. The average state is only 7.6 percent away from the frontier in the 98 percent LP model and, judging from the R^2 of 0.48 in equation (10), about half of that distance must be attributed to definitional problems in the variables rather than productive inefficiencies. An average loss of about 3–4 percent seems to be the most that is likely—a fairly small amount given the potential for quick expansion of U.S. agricultural output.

Are all states then equally good producers of agricultural output? Are there really no welfare losses due to poor decision making in U.S. agriculture? The answer to both questions is no. First, by measuring technical efficiency at the state level, all differences among farmers within the states were aggregated away. It is entirely possible for substantial productive inefficiencies to exist within states with few observed differences between states. The empirical results reported here can make no judgment at all on that score.

Second, the significance of allocative inefficiencies should not be ignored. Harberger (1954) has taught us that allocative inefficiencies have small welfare impact in a competitive setting, because the losses are only marginal rather than total, but parts of U.S. agriculture are largely outside the competitive equilibrating system. Land and most buildings are immobile. Farm labor can and does move out of agriculture, but—for a variety of economic and social reasons—it does so slowly, even in the face of substantial underemployment. Thus the losses to the more fixed factors are total and not just marginal because they are not employed elsewhere in the economy. The welfare costs of this type of allocative inefficiency are very great, particularly because the human resource is one that tends to be wasted.

References

Aigner, D. J., and Chu, S. F. "On Estimating the Industry Production Function." *A.E.R.* 58, no. 4 (September 1968): 826–39.

Brewster, John M., and Parsons, Howard L. "Can Prices Allocate Resources in American Agriculture?" *J. Farm Econ.* 28, no. 4 (November 1946): 938–60.

Comanor, William S., and Leibenstein, Harvey. "Allocative Efficiency, X-Efficiency, and the Measurement of Welfare Losses." *Economica* (August 1969), pp. 304–9.

Farrell, M. J. "The Measurement of Productive Efficiency." *J. Royal Statis. Soc.* (ser. A, General) 120, pt. 3 (1957): 253–81.

Griliches, Zvi. "Estimates of the Aggregate Agricultural Production Function from Cross-sectional Data." *J. Farm Econ.* 45, no. 2 (May 1963): 419–28. (*a*)

———. "The Sources of Measured Productivity Growth: United States Agriculture, 1940–60." *J.P.E.* 71, no. 4 (August 1963): 331–46. (*b*)

———. "Research Expenditures, Education, and the Aggregate Agricultural Production Function." *A.E.R.* 54, no. 6 (December 1964): 961–74.

Harberger, Arnold. "Monopoly and Resource Allocation." *A.E.R.* 44, no. 2 (May 1954): 77–92.

Hoch, Irving. "Estimation of Production Function Parameters Combining Time-Series and Cross-Section Data." *Econometrica* 30, no. 1 (January 1962): 34–53.

Leibenstein, Harvey. "Allocative Efficiency vs. 'X-Efficiency.'" *A.E.R.* 56, no. 3 (June 1966): 392–415.

Mundlak, Yair. "Empirical Production Function Free of Management Bias." *J. Farm Econ.* 43, no. 1 (February 1961): 44–56.

Timmer, C. P. "On Measuring Technical Efficiency." Ph.D. dissertation, Harvard Univ., 1969.

———. "On Measuring Technical Efficiency." *Food Research Institute Studies in Agricultural Economics, Trade, and Development* 9, no. 2 (1970): 99–171.

U.S. Department of Agriculture. Economic Research Service. *Farm Income: State Estimates, 1949–1967.* FIS-211, suppl., August 1968.

Walters, A. A. "Production and Cost Functions: An Econometric Survey." *Econometrica* 31, nos. 1–2 (January–April 1963): 1–66.

Yotopoulos, Pan A. "From Stock to Flow Capital Inputs for Agricultural Production Functions: A Microanalytic Approach." *J. Farm Econ.* 49, no. 2 (May 1967): 476–91.