Research Report CCS 308

MEASURING THE EFFICIENCY OF DECISION MAKING UNITS WITH SOME NEW PRODUCTION FUNCTIONS AND ESTIMATION METHODS

by

A. Charnes*
W. W. Cooper**
E. Rhodes***

August 1977

*The University of Texas, Austin, Texas.
**Graduate School of Business Administration, Harvard University.
***School of Urban and Public Affairs, Carnegie-Mellon University.

This paper represents a second revision of an earlier version entitled "Exposition, Interpretation, and Extensions of the Farrell Efficiency Measure." See [12] in the bibliography at the end of this paper.

This research was partly supported by NSF Grant No. SOC76-15876 "Collaborative Research on the Analytical Capabilities of a Goals Accounting System," and by Project NR047-021, ONR Contract Numbers N00014-75-C-0616 and N00014-75-C-0569 with the Center for Cybernetic Studies, The University of Texas, and ONR Contract N00014-76-C-0932 at Carnegie-Mellon University School of Urban and Public Affairs.

CENTER FOR CYBERNETIC STUDIES

A. Charnes, Director
Business-Economics Building, 203E
The University of Texas
Austin, Texas 78712
(512) 471-1821

ABSTRACT

A series of linear programming models are used to clarify and extend a measure of efficiency introduced by M. J. Farrell. The duals to these models are shown to yield estimates of production coefficients from the same empirical data and computations that yield the measures of efficiency. The nature of the resulting production functions and ways in which they differ from more customary ones are discussed en route to synthesizing the associated cost functions and other such (economic) relations. Methods for adjusting observations are suggested for economic inferences and policy applications. Multiple output-multiple input extensions are effected via a new definition of efficiency which involves a nonlinear model for determining the optimal input and output weights from observational data. The theory of fractional programming is used to secure ordinary linear programming models from which the weights and efficiency measures may also be obtained.

I. Introduction

Economic studies of resource allocation and related decision-making activities of governmental agencies and programs can encounter difficulty because of the form in which the underlying economic constructs are stated. Oriented mainly toward private markets and entities, the constructs from economic theory usually take the form of extremal relations. The concept of a production function which we shall be utilizing here is a case in point. This function assumes the form of an extremal relation from the property that output is always maximized from any combination of inputs that may be utilized.

When perfect competition prevails in relevant private sector markets, then considerations like the following may be relied upon: (1) All firms in the same market have approximately the same efficiency as the most efficient among them and (2) the latter firms, i.e., the relatively most efficient, are either at or near the maximum efficiency that the existing state of technology admits. The first, which may be referred to as relative efficiency, is the one we shall focus upon. The second almost inevitably carries with it such properties as freedom of exit and entry in pursuit of maximum profitability, etc., and hence it also carries with it possible flows from one "industry" to another, which are not characteristic of governmental (and like) activities which form the central concern for the formulations in this paper.

This paper is directed toward approaches to relative efficiency, e.g., for governmental activities, by means of which we can either (a) evaluate the efficiency of various decision-making entities participating in like programs or (b) adjust and allow for efficiency variations among these entities in order to bring the thus adjusted observations into conformity with underlying concepts involving extremal relations.

The latter topic leads to a technique of data analysis which we shall refer to as Data Envelopment Analysis,[1] but which we must reserve for separate treatment in order to deal adequately with the topics of the present paper. Hence we shall here only adumbrate this DEA approach. This is to say that our present paper will be mainly of a mathematical-economics variety which, for the most part, leaves aside the further topics of statistical-econometric data treatment. This is a natural way of proceeding, we might remark, since our DEA approach posits a prior adjustment of the data to bring them into conformance with the underlying extremal relations required by the constructs in economics before undertaking the related statistical hypotheses tests or estimates. Our present paper is also of value for empirical use in its own right, too, since it supplies new methods of effecting estimates of these extremal relations via mathematical (as distinguished from statistical) means which can be useful in a variety of contexts. For example, the methods in the paper may be used when one is satisfied that the data, (1) are not subject to "large" amounts of error, and (2) cover the relevant entities so that sampling of a statistical variety, say, is not a source of data variation.

That is, like in engineering and other sciences, we are supposing that the paradigms [2] of statistical estimation and testing need not be used when such errors are small relative to other considerations. In a similar vein, one further justification for our proposed approach arises from the fact that customary statistical econometric techniques based on "averaging,"

---

[1] See Rhodes [30].

[2] The term paradigm is used in the sense of Kuhn [26].

and like "central tendency" relations, are not wholly suited to the task of estimating extremal relations.[1] The estimating methods we are suggesting, on the other hand, are based on the duality relations of mathematical programming and hence on the extremal mathematics which underlies these relations. As such they are well suited to the estimation of extremal relations from empirical data provided we can reorient these mathematical programming constructs from their usual pre-decision uses as planning guides and thereby obtain a complementary use by applying these same methods and theorems to post-decision (empirical) data. Such, of course, is the main thrust of the present paper in which, we also note, we restrict ourselves to inequality systems so that we may utilize ordinal as well as cardinal arrays of data.

Having now highlighted some of our objectives in the present paper, we should pause to observe that empirical approaches to the study of production functions have not always relied on the extremal formulations of economic theory. Thus, R. A. Frisch, in what is probably the earliest study of production functions at the individual plant level recognized the difficulty involved in his use of a least-squares regression approach for estimating such extremal relations.[2] In particular, he explicitly noted the need for a production function formulation based on the "averaging" that such statistical estimating procedures imply, at least as a matter of practical application.

For purposes such as statistical prediction of actual behavior, etc., this kind of "practical empiricist" approach can have advantages. It should be recognized, however, that a reliance on such "averaging" methods produces

---

[1] See the discussions in [1], [2] and [3].

[2] An excellent discussion of Frisch's concept as well as other approaches to empirical uses production function studies may be found in [20].

estimates and relations which do not then have direct access to the large body of theorems and methods which are available from standard versions of micro-economic theory.[1] This is to say that departures from the defined extremal relations on which that theory rests require either separate and explicit justifications to utilize the related "marginalist" conventions and theories or else entirely new formulations must be essayed before anything like these conventions can be utilized.

In many ways the latter course is the one that will be followed in this paper. Although the production function concept that we shall be employing is not wholly new, having been employed, e.g., by M. Farrell [18] and others, it does differ from the ones that are customarily employed.[2] We also alter and extend even the production functions that Farrell et.al. utilized and we also supply new methods of estimation and interpretations.

En route to our fully developed formulation we shall try to stay fairly close to standard micro economic concepts. The point to be made for the present, however, is that our production functions will be based on optimizing models and methods. This means that we shall feel free to utilize the results of such optimizations to characterize possible economizing in resources associated with inputs and/or augmentations associated with outputs. We shall also utilize standard economic constructs such as marginal productivities, etc., and indeed we shall also utilize still other optimizations to derive still other economic relations such as the economic opportunity cost functions associated with our production functions.

All of the above will be formulated within the characterization we

---

[1] In many of the current energy studies the failure to make such adjustments seems especially in need of explicit justification if only because of the likely presence or "waste" (i.e., failure to achieve efficiency frontiers) in the presence of a large and rapid price change for a previously greatly "underpriced factor of production like energy. See [15].

[2] Actually neither Farrell nor his followers ever developed or employed these production function concepts per se.

gave for relative efficiency in our opening two paragraphs. This means that we will not be concerned with best alternative uses (e.g., in some other industry) but only with the release of resources, say, for which some other use is available. This will also all be accomplished within the limits of specified models and methods -- which will become clearer as we proceed. In any case our developments will be formulated around "decision-making entities" that we shall refer to as "firms." [1] Such "firms" will then be said to constitute an "industry" or to be members of a "program" when they have all of their inputs and outputs in common. With (1) all inputs and outputs in common at non-zero values we can obtain certain simplifications in dealing with issues like scaling to common units, etc.,[2] and (2) utilizing observations from individual firms permits us to avoid troublesome issues of aggregation, etc. Together these two approaches permit us to obtain added flexibility for synthesizing a variety of production functions which may then be aggregated or disaggregated to varying degrees in accordance with certain general prescriptions that we shall supply.

---

[1] These may also be different decision-making entities within a single agency such as, e.g., the activities under a single government department or program director.

[2] See the appendix to this article.

Finally, we shall make explicit contact with ongoing bodies of research in the following manner. First, we shall explicitly relate our developments to earlier work by M.J. Farrell and, indeed, the first part of our development will rest on this work.[1] After we have eliminated or modified some of Farrell's concepts,[2] however, we shall undertake certain new extensions. In particular we shall provide a linear programming formulation which via its dual automatically supplies the wanted production coefficient estimates that Farrell (and others) could not supply without a good deal of extra effort.[3] That is, this effort had to be undertaken in addition to the work required to secure the wanted estimates of decision-making efficiencies whereas our one linear programming approach secures both sets of estimates with the same computations.[4]

---

[1] See [12].

[2] E.g., Farrell's use of "points at infinity" to complete his analysis, as discussed in [18].

[3] Førsund [20] specifically notes the fact that Farrell fails to provide what is needed for a completely satisfactory characterization in these respects.

[4] Our model is also formulated in a way that provides easy access to standard linear programming codes. In this connection, however, we need to note that Alan Hoffman suggested a linear programming formulation at a very early date. See pp. 284-285 in [18]. In fact, the formulation which Hoffman presented as a discussant in [18], along with his suggestion for the use of Lemke's [29] "dual method" was subsequently employed with considerable gain in computing efficiency by Farrell and Fieldhouse in [19]. Although we, too, shall employ Lemke's "Dual Method" as well as Dantzig's [16] Simplex Method of solution - see Chapter XI in [9] for a combined geometric interpretation of both the Dual and Simplex Methods - we need to observe that our formulation differs from Hoffman's. His formulation, which was restricted to the single output case, is (a) presented in equality form and he does not deduce our inequality form and (b) carries with it such concepts as "points at infinity" by approximation with "large" real multiples of unit sectors. Moreover the uses of duality with related interpretations did not enter into Hoffman's discussion in any way.

With the above tasks completed we shall then undertake to make explicit contact with the Shephard-Samuelson[1] duality theory.[2] This will make available yet another body of developments in economics by means of which one can move from production functions to cost functions and related pricing and costing theories as well. Here we shall only carry the analysis far enough to make the indicated contact although we have elsewhere shown [15] that Shephard's "gauge function" is the complement of Farrell's efficiency measure -- and we have indicated other relations as well so that, e.g., an extension of "Shephard's lemma" is available for the kind of inequality analysis which is required for many types of governmental activities.

Such governmental activities frequently involve multiple outputs and multiple inputs with no evident weightings readily available via market prices, say, or other such objectively verifiable magnitudes. We may perhaps make the point at issue sufficiently clear by means of educational studies. If one wants to move from the usual scalarizations such as, e.g., income-earning ability as a single output from all educational inputs, then one confronts the need for weighting <u>cognitive</u> outputs such as "arithmetic ability" relative to <u>affective</u> outputs such as "positive attitude toward the community" relative to outputs of <u>psychomotor skills</u> such as

---

[1] See Shephard [36] and [37] and Samuelson [32].

[2] We have elsewhere argued that this is better regarded as a version of transform theory in mathematics not only to avoid confusion with the duality theory of mathematical programming but for other reasons as well. See [15].

"ability to tread water in a swimming pool and turn from front to back in a specified time." The latter, suitably measured as an index, say, of "bodily control and appreciation" can be regarded as an output which has "some value," but that value is not easily referred to any market system for a more or less objective determination of its relative weight. The same applies to "affective outputs" of the kind already indicated and also, to a lesser extent, to "cognitive outputs" such as arithmetic abilities.

Similar remarks may apply to the inputs. "Amount of teacher time" may be referred to salaries and perhaps further adjusted for inter-school and inter-program comparisons in a seemingly reasonable manner. But great difficulties may be encountered in dealing with inputs such as "the time and efforts of community leaders" and/or "parents, and others," which are evidently of value, but which have no ready market or other such referent for securing suitable sets of weights.

Since we want to deal with these multiple output-multiple input situations directly--i.e., without scalarizing in an ex-cathedra fashion--we shall proceed as follows. First we shall extend the customary definition of "efficiency" as a ratio of "wanted outputs" to "valued inputs" to formulate a new nonlinear programming problem. This formulation will enable us to determine the associated weights objectively and in an optimal manner from observational data. That is, the model together with its interpretation will make the nature of the optimization evident. The weight determination will then be objectively verifiable by anyone who applies this model to the same body of data and, allowing only for the case of alternate optima, a different choice of weights will require some other specified model in any case where the ratio definition of efficiency is applied.

The above model, and its interpretation, will provide conceptual clarity. Because the resulting problem is nonconvex as well as nonlinear, however, some additional developments will be needed for purposes of solution. Via the theory of fractional programming,[1] which we have elsewhere formulated, however, we shall be able to show how to transform this nonlinear and non-convex problem into an ordinary linear programming equivalent.

This transformation, as we shall see, also provides a natural generalization from the single output to the multiple output case. The linear programming problem resulting from this fractional programming transformation also has a dual.[2] This means that all of the preceding results for estimation and cost function relations, etc., go over in a readily apparent manner to this more general case. In other words, our new estimation methods for extremal relations go over to the estimation of multiple, as well as single, output relations.

Because these ideas involve a rather complex array of considerations, we have elected the following presentation strategy. Our initial development will be rather leisurely and conducted in some detail. The numerical illustrations as well as the mathematical developments and interpretations that we shall supply are intended to give the reader some "feel" for what is involved. To a considerable extent we shall also try to relate our discussion to familiar developments (like isoquant analysis) and already available literature

---

[1] This theory is complete as observed in [7] and [10], so that no difficulty need be encountered because of zeros in the denominators of the indicated ratios--or any other like aberrant behavior.

[2] I.e., a linear programming dual as distinguished from the dual to the corresponding fractional programming problem. See, e.g., Jaganathan [23] and Schaible [32]. See also Bector [5].

(such as Farrell). This done, however, we shall then accelerate our pace with only brief indications of possible uses for our more general developments. Then we shall draw all of the preceding together and try to point up possible further paths of research.

2. Background Illustration

Farrell in [18] employed three concepts of efficiency: (1) technical efficiency, (2) price efficiency and (3) overall efficiency (price and technical efficiency). Like Farrell, and others who have followed him, we propose to restrict ourselves mainly to his concept of technical efficiency.[1] Some of the reasons for this have already been indicated. In addition, price data are rather slippery and elusive in that they may impound a variety of motives relative to past, present and future prospects[2] and they may inject certain types of instability as these motives and their associated expectations change.

These and other problems associated with the use of market price data are discussed in detail by Farrell[3] as well as others.[4] Hence, we may dispense with further consideration of "price" and "overall efficiency" in order to focus our own attention on technical efficiency only. Here we need assume only that all inputs and outputs have value. That is, these inputs and outputs are not explicitly priced or costed but they all have value[5] in the sense that technical efficiency is not present when it is possible to increase any output without decreasing any other output or increasing any input. Conversely, technical

---

[1] To deal with problems such as increasing returns to scale, Farrell felt obliged to return to a use of "price" data. See [19]. See also Seitz [35].

[2] That is, market prices reflecting actual transactions as distinguished from the "duality price relations" which are really the same as optimal substitution ratios.

[3] See [18] and [19].

[4] Førsund in [20] provides a detailed discussion.

[5] This, as we shall see, is related to the concepts of Pareto-Koopmans' optimality as discussed in Chapter IX of [9].

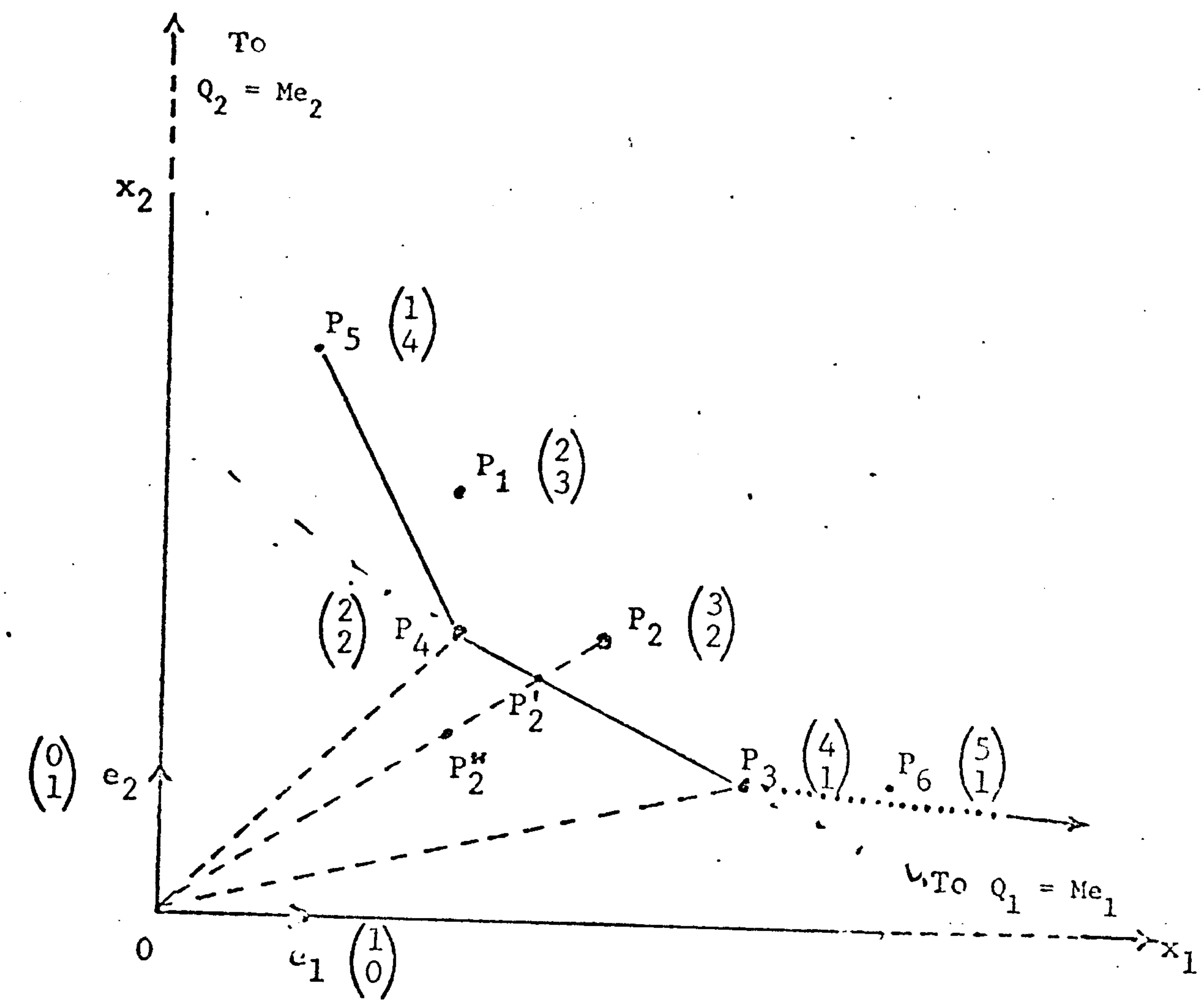
efficiency is also not present when it is possible to decrease any input without increasing any other input or decreasing any output. We would like to arrange our model and methods of measurement so that efficiency[1] will be ascribed to the observed behavior of any firm only if neither of the preceding is true.

Figure 1, below, provides an illustrative hypothetical example which we can use to illustrate Farrell's approach as well as our own. For this simple illustration we proceed as follows. First we reserve P_6 for separate discussion. Then we interpret $P_1, \ldots P_5$ as points in a two-dimensional Cartesian space in which the coordinates represent observed amounts of two different factors of production used to produce a single product by five different "firms."

We are here concerned with the one output, as distinguished from the multiple output, case. We do have multiple inputs (here two), however, and, given these observations, the first problem is to locate the efficiency frontiers or rather the subset of extreme points from which such a frontier may be formed by reference to the five points, P_1, P_2, P_3, P_4, and P_5 as shown. These "efficient" points are then joined by linear segments, as in Figure 1, to form the corresponding "efficient isoquant" which, as can be seen, is a "piecewise linear" and continuous curve. Here in this two-dimensional space the efficient surface is a line, i.e., an isoquant line, but in m-dimensions, "facets" of dimension greater than one will be subtended in an efficient isoquant surface

[2]Actually a slight generalization of the idea of Pareto-Koopmans' optimality in the sense of [9] Chapter IX is involved here in that we are ascribing "value" to inputs as well as outputs.

Figure 1


To
$Q_2 = Me_2$
x_2
P_5 $\binom{1}{4}$
P_1 $\binom{2}{3}$
$\binom{2}{2}$ P_4
P_2 $\binom{3}{2}$
P_2'
P_2''
P_3 $\binom{4}{1}$
P_6 $\binom{5}{1}$
$\binom{0}{1}$ e_2
To $Q_1 = Me_1$
0
e_1 $\binom{1}{0}$
x_1

An "efficient isoquant" in this analysis represents what Farrell regards as "practically attainable efficiency." By this he means that at least some "firms" attained this level of efficiency even though others did not. Then, by interpolation between these observed points, he provides a "facet" (here a line) of reference for measuring the relative efficiency of the remaining firms.

To develop his measure of efficiency, Farrell makes two key assumptions. The first is that any firm can expand or contract along a "ray" through the origin such as the one indicated by the broken line from the origin to P_2. The other assumption is that all convex combinations of the observed points, $P_1, \ldots, P_5$ represent actual production possibilities. The "efficient points" then correspond to the extreme points of this convex set plus the union of the boundary facets of this convex set, i.e., the union of the boundary simplexes formed on the efficient extreme points. The latter thus constitutes the "efficient isoquant" or the "efficient facet" - *viz.*, the locus of points corresponding to the minimum inputs of all factors for a specified level (e.g., one unit) of the observed output.

To see what is involved graphically in Farrell's proposed measure refer to Figure 1 where the "efficient isoquant" is drawn in piecewise linear segments from P_5 to P_4 and P_4 to P_3. These points constitute the input vectors, per unit output, for firms 3, 4, and 5 with linear interpolation specifying what should be attainable for P_1 and P_2. For example, let

$\ell(OP_2)$ = length of "ray from origin" to P_2

(1) and

$\ell(OP_2')$ = length of ray from origin to P_2' which is the efficient isoquant point "closest to the origin" on the ray to P_2.

Then Farrell's measure of efficiency or, more precisely, his measure of "technical efficiency" is

(2) $0 \leq TEF(P_2) = \ell(OP_2')/\ell(OP_2) \leq 1.$

Note that if P_2' had actually occurred its measure of efficiency would have been unity. This is to say that because P_2' is a point on the boundary line segment between P_4 and P_3 it is an efficient isoquant point and therefore the "closest to the origin" isoquant point on the ray to P_2'. We note that since P_3 and P_4 are adjacent efficient extreme points of the convex set of production possibilities, P_2' can be represented as a convex combination of these points. That is, we would have $P_2' = P_3\lambda_3' + P_4\lambda_4'$ with $\lambda_4', \lambda_3' \geq 0$ and $\lambda_3' + \lambda_4' = 1$. Since, in fact, P_2 is observed <u>above</u> the line connecting P_4 and P_3, its efficiency measure is less than unity, which we shall subsequently relate to the expression $P_2 = P_3\lambda_3' + P_4\lambda_4'$ with $\lambda_3', \lambda_4' \geq 0$ and $\lambda_3' + \lambda_4' > 1$. Moving next to P_2'' we can also express this as $P_2'' = P_3\lambda_3' + P_4\lambda_4'$, with $\lambda_3', \lambda_4' \geq 0$, $\lambda_3' + \lambda_4' < 1$. In each case non-negativity of both variables is preserved with the resulting sum indicating where the point is relative to the isoquant segment connecting P_3 and P_4.

Now consider non-negativity. All points inside the cone with edges indicated by the broken line from the origin through P_3 and P_4, respectively, can be expressed as nonnegative combinations of these two points. Only when one attempts express a point outside the cone, e.g., a point such as P_1, is the nonnegativity condition violated. But then a recourse to our piecewise linearity assumption allows us to generate a new isoquant segment connecting P_4 and P_5 from which to evaluate the relative efficiency of P_1, and so on. Clearly, then, any point such as P_1 or P_2 will <u>not</u> be efficient because there exists a convex combination of other points which has a lower value for at least one of its input coordinates. This is not the case, however, for the other points P_3, P_4, P_5 in the set we are presently considering or for any point lying on the lines generated from all convex combinations of P_5, P_4 and P_4, P_3, respectively.

Thus, relative to the indicated set of observations, we characterize only the latter set of points as efficient. The others are not efficient on Farrell's "ray" and "isoquant" assumptions. Bear in mind, however, that we are directing our attention solely to relative evaluations and technical efficiency only.

Having achieved our wanted clarity in only qualitative (efficiency vs. inefficiency) form, however, we now need to provide a way of meaningfully measuring the "amount" or "degree of efficiency" displayed by each of our observed firms. We also want to do this is a way that provides a convenient basis for subsequent generalizations. Therefore we proceed as follows. For each of j = 1, ..., n firms we are given observations on their inputs for each of two factors of production which we symbolize as

$$\begin{aligned} x_{1j} &= \text{amount of first factor used by firm } j \\ x_{2j} &= \text{amount of second factor used by firm } j \end{aligned} \tag{3.1}$$

We are also given the amount of output -- sales, shipments, or production -- which we symbolize as

$$(3.2) \qquad y_j = \text{amount of output of firm } j .$$

In each case we assume that the amounts in (3.1) and (3.2) are all positive and then combine them for ratio representation, as in Figure 1, via

$$(4) \qquad \begin{aligned} x'_{1j} &= x_{1j}/y_j \\ x'_{2j} &= x_{2j}/y_j \end{aligned}$$

In other words, x'_{1j}, x'_{2j} are, respectively, the inputs per unit output utilized by firm $j = 1, \ldots, n$. The points P_j in Figure 1 are then represented as

$$(5) \qquad P_j = \begin{bmatrix} x'_{1j} \\ x'_{2j} \end{bmatrix} .$$

Thus the observations in Figure 1 are all "normed" by reference to the outputs for, respectively, each of the $j = 1, \ldots, 5$ firms we are presently considering. Accordingly, the "efficient isoquant" exhibited there represents the required input rates for 1 unit of output and so we also refer to it as the "unit isoquant."

We can now conclude this section by further formalizing some of our preceding discussion as follows. Consider any two points P_r and P_s which form a basis for the 2-dimensional space exhibited in Figure 1. That is, we assume P_r and P_s, which are 2-component vectors, are linearly independent. Any other observed point, P_o, say, may be expressed in terms of them via

$$(6.1) \qquad P_r\lambda'_r + P_s\lambda'_s = P_o.$$

Now assume P_r and P_s are also efficient. If $\lambda'_r, \lambda'_s \geq 0$ then P_o is in the cone with edges formed by extending a ray from the origin through each of P_r and P_s, respectively. Conversely, if P_o is not in that cone then at least one of the variables required for the expression of P_o in (6.1) must be negative. If

(6.2) $\lambda'_r, \lambda'_s \geq 0$

and

(6.3) $\lambda'_r + \lambda'_s = 1$

then P_o is on the line segment connecting P_r and P_s. If, on the other hand, $\lambda'_r + \lambda'_s > 1$ then P_o cannot be efficient while if $\lambda'_r + \lambda'_s < 1$ then P_r and P_s cannot both be efficient.

In any case since we are considering points on a ray through the origin, their coordinates in terms of any basis are proportional. Thus we have

(6.4) $0 \leq \text{TEF}(P_o) = 1/(\lambda'_r + \lambda'_s) \leq 1$

with (6.2) and (6.3) also applying whenever P_r, P_s and P_o are all efficient and in the cone. Alternatively, if <u>only</u> condition (6.2) applies, then at least one of P_r, P_p and P_o is not efficient while if (6.2) does <u>not</u> apply then these three points are not in the same cone, and hence cannot be used to evaluate each other's relative (technical) efficiency.

Our subsequent linear programming formulations and developments are undertaken with this kind of piecewise linear development in mind. Hence, we pause for remarks like the following. Translated into economic terms this means that the marginal rates of substitution are piecewise constant along the indicated efficiency frontier. Moreover, these substitution constants, which vary from one piecewise segment to another, do not affect the marginal productivities of any other pair of factors, which are also piecewise constant (in their ratios) along their own respective frontiers.

Now we should observe that the above analysis obviously extends to other functional representations which can also be transformed into suitable piecewise linear format. A case in point would be a development via piecewise Cobb-Douglas functions. These functions being representable as linear under a logarithmic transformation were, in fact, utilized by Farrell in his original empirical studies [18]. These, and other like possibilities which may be present, will also be preserved in the development that follows.[1]

On the other hand, we will not discuss cases which are not transformable to such piecewise linear equivalents. Neither will we discuss other cases such as increasing returns to scale,[2] etc., which violate convexity (even after suitable transformations) and the related assumptions that we are also making. Finally, we may observe that Farrell's ray assumption imputes the property of constant returns to scale to each firm. Actually, this assumption is stronger than necessary since we only require that an intersection be effected with the convex isoquant at the indicated point, but to relax this ray assumption would involve introducing additional complexities into the analysis. Hence we continue with Farrell's ray assumption while observing that the assumed scale constant it implies will, in general, differ for every firm.

---

[1]Actually Farrell did not use piecewise Cobb-Douglas functions but, as Duffin, Peterson and Zener [17] pp. 265-267 show, the extension involved in optimizing a Cobb-Douglas function under Cobb-Douglas type constraints goes over into an ordinary linear programming problem. See also Charnes, Cooper and Kortanek [11] linear programming problem. See also Charnes, Cooper and Kortanek [11] for more general functions which can be represented in infinite systems of linear inequalities under suitable transformations.

[2]See Farrell and Fieldhouse [19] and Seitz [35].

3. Linear Programming Formulations

Now we turn to our proposed linear programming development in which we consider the case where there are n firms and m inputs so that

$$(7.1) \qquad P_j = \begin{bmatrix} x'_{1j} \\ \cdot \\ x'_{ij} \\ \vdots \\ x'_{mj} \end{bmatrix}$$

where

$$(7.2) \qquad \begin{aligned} x'_{ij} &= x_{ij}/y_j \\ i &= 1, .., m \\ j &= 1, .., n, \end{aligned}$$

which means that each of these input amounts is normed on y_j, the observed output for the jth firm, as before.

P_o is assumed to be a vector of observations associated with one of j = 1, ..., n firms whose inputs, per normed unit output, have been observed. For concreteness let $P_o = P_k$ and let it be desired to determine whether P_k is efficient. To do this we form the linear programming problem,

$$\text{(7.3)} \qquad \max \; z_o = \sum_{j=1}^{n} \lambda_j'$$

$$\text{with} \qquad \sum_{j=1}^{n} P_j \lambda_j' = P_o, \; \lambda_j' \geq 0$$

and, to insure that we consider P_o relative to a cone on efficient extreme points which contains P_o, we employ an adjacent extreme point algorithm such as the "simplex" or "dual" method to solve this linear programming problem.

We then have as one possible solution

$$\text{(8)} \qquad P_k \lambda_k' = P_k = P_o$$

with $\lambda_k' = 1$. It follows that Max $z_o = z_o^* \geq 1$ and therefore

$$\text{(9)} \qquad 0 \leq TE'(P_o) = 1/(\sum_{j=1}^{n} \lambda_j^{*\prime}) = 1/z_o^* \leq 1$$

as required.

We need to insure that the efficient isoquant, when attained, accommodates all possibilities. For instance, we need to insure that some firm using an extremely large amount of one factor and a very small amount of every other factor will still be on a ray which intersects the efficient isoquant en route to the origin. Farrell adds "points at infinity" to insure this but this is awkward and so we now undertake a development which enables us to move ahead without recourse to these constructs.

We can represent the points at infinity as $Q_i = Me_i$, $i = 1, \ldots, m$, where M is a non-Archimedean transfinite number ($M > r$, r any real number),[1/] and the e_i are the unit coordinate vectors - see, e.g., e_1 and e_2 in Figure 1. The other points P_j have finite and non-negative entries. But now consider the non-negative representation of P_o, a finite, non-negative vector, in terms of the P_j and Q_i - e.g.,

$$P_o = \sum_{j=1}^{n} P_j \lambda'_j + \sum_{i=1}^{m} Q_i \xi_i \tag{10.1}$$
$$= \sum_{j=1}^{n} P_j \lambda'_j + M \sum_{i=1}^{m} e_i \xi_i$$

where

$$\lambda'_j, \xi_i \geq 0 .$$

Since the entries in the P_j and e_i are finite and non-negative, such a representation is possible in terms of finite λ'_j if and only if the ξ_i are of the form

$$\xi_i = \frac{1}{M} s_i \tag{10.2}$$

where the s_i are finite and non-negative. Thus, we get the equivalent finite representation of the constraints as

$$\sum_{j=1}^{n} P_j \lambda'_j + \sum_{i=1}^{m} e_i s_i = P_o \tag{10.3}$$
$$\lambda'_j, s_i \geq 0 .$$

To complete this part of the analysis we now explicitly formulate the linear programming problem with points at infinity as follows,

1/ See, e.g., Robinson [31].

$$\text{Max} \quad z_o = \sum_{j=1}^{n} \lambda_j' + \sum_{i=1}^{m} \xi_i$$

(11.1)
$$\text{with } P_o = \sum_{j=1}^{n} P_j \lambda_j' + \sum_{i=1}^{m} Q_i \xi_i$$

$$\lambda_j', \ \xi_i \geq 0.$$

Then we replace this with the following new form

$$\text{Max} \quad z_o = \sum_{j=1}^{n} \lambda_j' + \frac{1}{M} \sum_{i=1}^{m} s_i$$

(11.2)
$$\text{with } P_o = \sum_{j=1}^{n} P_j \lambda_j' + \sum_{i=1}^{m} e_i s_i$$

$$\lambda_j', \ s_i \geq 0 ,$$

where we also note that the constraints can be written in our newly established equivalent inequality form as

(12)
$$P_o \geq \sum_{j=1}^{n} P_j \lambda_j'$$

$$\lambda_j' \geq 0, \ j = 1, \ \ldots, \ n.$$

This inequality format for the constraints can be given a geometric interpretation in terms of relative efficiency of two production possibilities - namely, P_i is at least as efficient as P_j if $P_i \leq P_j$. In this manner the Farrell efficient (or frontier) points are simply "vector" or "Pareto minimal" points of the set of all production possibilities. This means that P_j is not efficient relative to P_i if P_i contains at least one entry strictly less than the corresponding entry in P_j.

A further geometric interpretation of the set of production possibilities can also be made. It is a convex set derived from the convex hull of the observed points by adding to this convex hull all points northeast of it. Thus, if P_r is a production possibility then so is P_s if $P_s \geq P_r$.

The non-Archimedean programming problem (11.2) provides us with a z_o^* whose reciprocal is the Farrell efficiency measure up to a (non-Archimedean) infinitesimal. If any of the s_i - which may be considered as slack variables for the representation (12) - is in at positive value in an optimal solution, then P_o cannot be efficient. If all s_i are at zero value, then the infinitesimal part of z_o^* is zero. Thus, for our new computational version of efficiency we can replace our earlier convexity characterization with

(13) $P_k = P_o$ is efficient if and only if

i. Its optimal z_o value is $z_o^* = 1$, and

ii. The slack variables are all at zero value in every optimum tableau,

for the problem

(14.1)
$$\text{Max} \quad z_o = \sum_{j=1}^{n} \lambda'_j$$
with
$$\sum_j P_j \lambda'_j \leq P_o$$
$$\lambda'_j \geq 0$$

or its equivalent

(14.2)
$$\text{Max} \quad z_o = \sum_{j=1} \lambda'_j$$
with
$$\sum_{j=1}^{n} P_j \lambda'_j + \sum_{i=1}^{m} e_i s_i = P_o$$
$$\lambda'_j, \ s_i \geq 0 \ .$$

4. Computational Example

We illustrate the above developments via the data of Figure 1. To initiate this illustration we shall continue to consider only the points P_1, ..., P_5 and reserve P_6 for separate consideration. This leads to the tabular arrangement in Table 1 where, with the conventions of [9], we use B and c_B to designate basis vectors and their functional coefficients at each stage. Thus, at "Stage 0" the basis B consists of the slack vectors e_1 and e_2 with functional coefficients of zero shown explicitly on the left.

(Insert Table 1)

Using full (rather than contracted) tableau arrangements for effecting our simplex calculations, we proceed from "Stage 0" to an optimum at "Stage 3*" from which we secure the following information: P_4 and P_5 are in the optimal basis and hence are both efficient. Concomitantly, $P_o = P_1$ has $z_o^* = 7/6 > 1$ and hence is not efficient - see i in (13) - and, in fact,

$$\text{TEF}(P_o = P_1) = 1/z_o^* = 6/7$$

measures the reduction in factor inputs which could have put this firm on the unit isoquant if it had been producing efficiently - *viz*.,

TABLE 1

Illustration of Simplex and Dual Method Calculations

SIMPLEX METHOD

Stage	Basis		Structural Vectors					Slack Vectors		Stipulations Vector	Optimal Basis
	c_B	B	P_1	P_2	P_3	P_4	P_5	e_1	e_2	$P_o = P_1$	
0	0	e_1	2	3	4	2	1	1		2	
	0	e_2	3	2	1	2	(4)		1	3	
	$z_j - c_j$		-1	-1	-1	-1	-1				
1	0	e_1	5/4	5/2	(15/4)	3/2		1	-1/4	5/4	
	→ 1	P_5	3/4	1/2	1/4	1/2	1		1/4	3/4	
	$z_j - c_j$		-1/4	-1/2	-3/4	-1/2			1/4	3/4	
2	→ 1	P_3	1/3	2/3	1	(2/5)		4/15	-1/15	1/3	
	1	P_5	2/3	1/3		2/5	1	-1/15	4/15	2/3	
	$z_j - c_j$					-1/5		1/5	1/5	1	
3*	→ 1	P_4	5/6	5/3	5/2	1		2/3	-1/6	5/6	P_4, P_5
	1	P_5	2/6	-1/3	-1		1	-1/3	2/6	2/6	
	$z_j - c_j$		1/6	1/3	1/2			1/3	1/6	7/6	

DUAL METHOD

Stage	c_B	B	P_1	P_2	P_3	P_4	P_5	e_1	e_2	$P_o = P_2$	Optimal Basis
4	1	P_4	5/6	5/3	5/2	1		2/3	-1/6	5/3	
	← 1	P_5	2/6	-1/3	(-1)		1	-1/3	2/6	-1/3	
	$z_j - c_j$		1/6	1/3	1/2			1/3	1/6	4/3	
5*	1	P_4	5/3	5/6		1	5/2	-1/6	2/3	5/6	P_4, P_3
	1	P_3	-1/3	2/6	1		-1	2/6	-1/3	1/3	
	$z_j - c_j$		1/3	1/6			1/2	1/6	1/3	7/6	

$$6/7P_1 = 6/7\begin{bmatrix}2\\3\end{bmatrix} = \begin{bmatrix}12/7\\18/7\end{bmatrix}.$$

That is, the operational meaning that we (like Farrell) accord to this measure is that a reduction of both of these inputs by the indicated ratio would have resulted in the same output if this firm had produced efficiently. Thus only with $TEF(P_o) = 1$ would no such reduction be possible so that the firm associated with P_o could be characterized as efficient.

Since P_4 and P_5 are efficient, there is a possibility that the basis of Stage 3* could also be used to measure the efficiency of other points besides P_1. In fact, this would be the case if any of the other vectors, P_2 and P_3, had only non-negative entries in their columns at this tableau stage, i.e., at Tableau Stage 3*. However, this is not the case. Both P_2 and P_3 have negative entries in their columns -- which means that neither they, nor the rays to them (from the origin), lie on or intersect the line segment connecting P_4 and P_5. In other words, P_3 and P_2 lie outside the cone generated from the origin through P_4 and P_5 and hence cannot be expressed as non-negative combinations of these two points. Cf. Figure 1 and the discussion in Section 2.

We next observe, with A. J. Hoffman in [18] that we are now in a position to use C. E. Lemke's "dual method".1/ This, too, is illustrated in Table 1 by according P_2 the status of a new P_o and transferring the data for it at Stage 3* to the new tableau represented as "Stage 4". We effect this transfer only for clarity, however, since we could also have proceeded directly from Stage 3* without any tableau such as Stage 4 to achieve the new optimum shown at Stage 5*. This is all done via the dual method, as indicated, and hence there is no need to start afresh or even to backtrack to earlier tableaus.

1/ See [29]. For further treatment of the dual method also see Chapter XI of [9].

The results in Stage 5* show that P_4 and P_3 are efficient and that P_2 is <u>not</u> efficient. Indeed, P_2 is "exactly" as inefficient as P_1. Observe, however, that P_1 was referred to P_4 and P_5 while P_2 was referred to P_3 and P_4, so that the efficient referents are not the same. The term "exactly" must, of course, be taken with a "grain of salt" in any case, since there is some room for observational error which would presumably be evident on replication,[1] and this is a weakness which we referred to in our introduction.[2]

On the other hand, the fact that the referents generally differ for each firm which is to be evaluated can have advantages for certain kinds of applications. In evaluating educational programs, for instance, the factor inputs utilized by different "firms" (e.g., schools or school districts) may differ from one area to another because of legal or other requirements. Hence we would li to have the basis for efficiency evaluations selected from among the efficient vect which are as much alike as possible to the vector being evaluated. For instance, we would prefer to have P_1 evaluated with P_5 and P_4 as referents and P_2 evaluated with P_4 and P_3 as referents instead of having P_1 and P_2 both evaluated with P_5 and P_3 as a basis. Loosely speaking, our model (and computational routines) have this property and, in addition, the vectors in the basis which are closest to the point being evaluated will also tend to receive the greatest weight.

---

1 The reader should also refer to the exchange between M. G. Kendall, M. Quenouille, and M. J. Farrell in [18].

2 See the discussion of our DEA approach to data treatment in [30].

5. Slack and Alternate Optima

The preceding remarks need some amplification which we shall try to supply in a way that also clarifies condition ii in (13). For this we now turn to

$$P_6 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}.$$

As depicted in Figure 1, this point lies to the right of the efficiency frontier which terminates at P_3. Via the developments provided in (10.1) ff., we can, of course, exhibit the nonefficiency of P_6 within the tableau for (14.2). We proceed, however, to a separate treatment of P_6 to show, inter alia, how some of the additional information available in these extended tableaus may be utilized.

First we observe that the inverse of any basis is found under the slack vectors at any stage.[1] Thus, continuing with (14.2) we immediately have

$$B^{-1} = (P_4, P_3)^{-1} = \begin{bmatrix} -1/6 & 2/3 \\ 2/6 & -1/3 \end{bmatrix}$$

from Stage 5* in Table 1. Hence, we also have

$$B^{-1}P_6 = \begin{bmatrix} -1/6 \\ 8/6 \end{bmatrix}$$

for insertion in the "P_o" column at this stage. As was the case at Stage 4, one of the components is negative, but we are, nevertheless, in a position to continue with the dual method, just as before.[2]

---

[1] See [9] and [16] for further discussion.

[2] Since P_k, the new vector to be considered, has $P_k \geq P_r$, the usual condition for continuation with the dual method is automatically satisfied. I.e., one has $z_j - c_j \geq 0$, all j, including j=k.

One iteration with the dual method produces

$$P_3 \lambda_3^{*'} + e_1 s_1^* = P_o = P_6$$

with

$$\lambda_3^{*'} = 1 \text{ and } s_1^* = 1$$

and all other variables at zero value. Evidently the zero slack requirement in ii of (13) is violated and so P_6 is not efficient. The interpretation is also relatively straightforward - viz., a reduction of $x_1 = 5$ to $x_1 = 4$ would eliminate this positive slack and bring P_6 into coincidence with P_3, in which case $\lambda_3^{*'} = z_o^* = 1$ and both conditions in (13) being satisfied efficiency would again be achieved.

Now we observe that an alternate optimum is present under (14.1) since also $z_o^* = 1$ with $\lambda_6^{*'} = 1$ and $s_1^* = 0$. However, as we have just seen, P_6 is not efficient relative to the basis $B = (P_3, e_1)$ since evidently a reduction in x_1 is possible relative to this basis.

Thus it is necessary to insure that no such slack is present in any alternate optimum. Among these alternate optima one would then select the one with maximal $\sum_{i=1}^{m} s_i^*$ which is equivalent to working with (11.2) rather than (14.2). In the present case this would give $z_o^* = 1 + \frac{1}{M} > 1$. In other words when we employ (11.2) rather than (14.2) we need utilize only condition i in (13). On the other hand, we can continue to work with (14.1) provided we are also willing to apply ii of (13) in the indicated manner since the only way that this ambiguity can occur is via the presence of positive slack in some alternate optimum.[1]

---

[1]This ambiguity, too, may be repaired in a more formal manner. We shall not pause to do this here, however, since in Section 8, below, we shall provide a more symmetric formulation which will reduce only to condition i of (13) so that condition ii of (13) may then be ignored.

6. Duality and Other Tableau Characterizations

We should stress the importance of employing an adjacent extreme point method in the calculations and characterization of extreme efficient points. In particular, as will be seen, the tableaus resulting from simplex and dual method calculations - which are both adjacent extreme point methods - provide a wealth of information concerning such things as efficient frontier facets and the normal vectors to them.

In order to map out the efficient isoquant we proceed by determining the extreme points spanning one facet plus the normal to this facet. We then move on through adjacent efficient facets and continue doing so until all of the wanted efficient facets and normal directions have been determined. The latter, which correspond to an optimal solution to the dual problem associated with the primal problem formed from (14.1) or (14.2), provides us with estimates of the isoquant slopes, as we shall show.

Our procedure allows us to characterize these efficient facets explicitly and in ways which have not heretofore been available. Hence, we shall develop them here in all detail and generality as follows.

We shall first assume that all of the observed points P_j have positive components. Also, if m is the number of such components, then we shall also assume that there is at least one basis for m-dimensional space among the vectors P_j. We can then start by choosing P_o as some point which is properly interior to the cone with apex at the origin and spanned by $P_1, \ldots, P_n$. This can always be arranged since, for example, $P_o = \sum_{j=1}^{n} P_j$ will have this property.

An illustrative example may help to fix ideas and provide concrete insight for our procedures and interpretations. Thus, going back to the data of Figure 1 (without P_6) and choosing $P_o = P_1$ initially, as before, we set up the simplex tableau for (14.2) and then proceed as in Table 1.

Notice that for (14.2) we always have a basis of natural slack unit vectors as in Stage 0. Notice further that the tableau entries under the structural vectors $P_1, \ldots, P_5$ and the slack vectors e_1, e_2 are completely independent of the entries under the stipulations vector P_o - i.e., the entries under the structural and slack vectors in each tableau are the same no matter what P_o is selected.

In particular, consider Stage 2 in Table 1. As is evident from Figure 1, both P_3 and P_5 are efficient extreme points. Nevertheless, the facet spanned by them is not efficient. Correspondingly, the Stage 2 tableau is not an optimal tableau since $z_4 - c_4 = -\frac{1}{5} < 0$. But Stage 3* is an optimal tableau even though $z_o^* = 7/6 > 1$ tells us that P_1 is not efficient. Now replacing P_1 by $P_o = P_4\nu_4 + P_5\nu_5$ for <u>any</u> $\nu_4, \nu_5 \geq 0$ and $\nu_4 + \nu_5 = 1$ will not change the $z_j - c_j$ but will give us a P_o column of ν_4, ν_5 with $z_o^* = 1$ so the whole facet spanned by P_4, P_5 is efficient.

The lesson here is that we can start with P_o an interior point and proceed through simplex tableaus to an optimal tableau. Then replacing P_o by any convex combination of the optimal basis vectors we will achieve a $z_o^* = 1$. Hence we will have exhibited an efficient facet and, as we shall show, the equation of the hyperplane containing this facet will also then be available. For, as is well known,[1/] the $z_j - c_j$ under the slack vectors in an optimal tableau <u>are</u> an optimal solution to the dual problem. I.e., these optimal $z_j - c_j$ values under the slack columns in the primal tableau are the optimal values $\omega = \omega^*$ for the dual problem,

(15)
$$\begin{aligned} &\text{Min } g_o = \omega^T P_o \\ &\text{with } \omega^T P_j \geq 1, \; j = 1, \ldots, n \\ &\omega^T \geq 0 \end{aligned}$$

[1/] See, e.g., [9] or [16].

and $\omega^{*T} P_i = 1$ for each P_i in an optimal basis. The superscript T represents transposition, as usual, so that, e.g., ω^T represents the transpose of the column vector ω with components $\omega_1, \ldots, \omega_m$ and ω^{*T} denotes an optimum value for these variables in the above problem.

To show that ω^{*T} is orthogonal to the efficient facet spanned by these P_i, we need only show ω^* is orthogonal to any direction lying in the facet, e.g., to any vector which is the difference, $\bar{P} - \bar{\bar{P}}$, of two vectors in the facet. Since $\bar{P}$ and $\bar{\bar{P}}$ are in the facet,

$$\bar{P} = \sum_i P_i \bar{\nu}_i, \quad \bar{\bar{P}} = \sum_i P_i \bar{\bar{\nu}}_i$$

$$0 \leq \bar{\nu}_i, \bar{\bar{\nu}}_i \leq 1$$

(16.1) and

$$\sum_i \bar{\nu}_i = \sum_i \bar{\bar{\nu}}_i = 1$$

where summation is over the indexes indicated by these P_i. But then

$$\omega^{*T} (\bar{P} - \bar{\bar{P}}) = \sum_i (\omega^{*T} P_i \bar{\nu}_i - \omega^{*T} P_i \bar{\bar{\nu}}_i)$$

(16.2)
$$= \sum_i (1\bar{\nu}_i - 1\bar{\bar{\nu}}_i) = \sum_i \bar{\nu}_i - \sum_i \bar{\bar{\nu}}_i = 0,$$

since $\sum_i \bar{\nu}_i = \sum_i \bar{\bar{\nu}}_i = 1$. Q.E.D. Hence, the ω^* corresponding to this efficient facet (and simplex) determined by this optimal basis is orthogonal (or normal) to it. Thus, ω^* is normal to the hyperplane containing this facet. The equation of this hyperplane is

(17) $$\omega^{*T} x = 1$$

where x is any point in the linear space spanned by the totality of the P_j's and e_i's.

We now return to the two-dimensional example of Figure 1 in order to exemplify these developments. Taking Stage 3* in Table 1 for illustration, we have $\omega^{*T} = (\frac{1}{3}, \frac{1}{6})$ as is apparent from the $z_j - c_j$ values listed under the slack vectors at this stage. The points P_o of the efficient facet are $P_o = P_4 \nu_4 + P_5 \nu_5$ for all $\nu_4, \nu_5 \geq 0$ with $\nu_4 + \nu_5 = 1$. Thereby we have $\omega^{*T} P_o = (\omega^{*T} P_4) \nu_4 + (\omega^{*T} P_5) \nu_5$.

Moving to (17) we therefore have, for this two-dimensional case,

$$\omega^{*T} x = \omega_1^* x_1 + \omega_2^* x_2 = 1$$

which, via the results available at Stage 3* in Table 1, become

$$\frac{1}{3} x_1 + \frac{1}{6} x_2 = 1.$$

In Figure 1, this facet is on a line with the segment connecting P_4 and P_5 constituting part of the efficient isoquant and, in fact, this segment may now be rendered explicitly as the set of points represented in

$$S\left[4,5\right] \equiv \left\{ (x_1, x_2) : \frac{1}{3} x_1 + \frac{1}{6} x_2 = 1; \quad 1 \leq x_1 \leq 2, \ 2 \leq x_2 \leq 4 \right\}.$$

Continuing to Stage 5* and observing the $z_j - c_j$ values under the slack at this stage we then obtain the expression for the adjacent isoquant segment as

$$S\left[3,4\right] \equiv \left\{ (x_1, x_2) : \frac{1}{6} x_1 + \frac{1}{3} x_2 = 1; \ 2 \leq x_1 \leq 4, \ 1 \leq x_2 \leq 2 \right\}.$$

Of course, the property to be emphasized is that the numerical values of the coefficients in these expressions are available without any extra effort.[1] Hence, also without extra effort we have a new method of effecting numerical estimates of extremal relations from observational data. This should be of interest for economics, or any other discipline where explicit numerical estimates of such relations are wanted -- e.g., under the kind of conditions remarked upon in our introduction.

7. Production Functions and Cost Relations

We now leave aside other issues like invariance of our efficiency measures[2] in order to concentrate on the production functions associated with these coefficient estimates. These functions are evidently not of the variety which have customarily been used in empirical studies. They are not of the aggregate variety which generally assume that (i) all firms

---

[1]The isoquant segment extending from P_3 to Q_1 may be represented by

$$S[3, 1') \equiv \{(x_1, x_2) : 1 = 0\ x_1 + x_2;\ 4 \leq x_1,\ 1 \leq x_2\}$$

or by

$$S[3, 1') \equiv \{(x_1, x_2) : 1 = \frac{1}{M} x_1 + x_2;\ 4 \leq s_1,\ 1 \leq x_2\}$$

where the prime refers to the subscript on Q and the square bracket denotes inclusion of the point whose subscript is next to the bracket, while the round bracket denotes exclusion. The expressions refer to the segments from P_3 to Q_1 in Figure 1, according to whether (11.1) or (14.2) is employed - and a similar development may obviously be undertaken for the segment extending from P_5 to Q_2.

[2]See Appendix.

have the same production function[1] and (ii) are on their efficiency frontiers as a necessary condition for obtaining access to the relevant micro theorems. Neither of these assumptions is used here[2] but, of course, we proceed on the assumption that relative, rather than theoretically attainable,[3] efficiency suffices for access to the theorems and procedures of micro analysis.[4] Indeed it is one purpose of the preceding development to distinguish between inefficient and efficient firms in such a way that only the latter enter into the coefficient estimates for these production functions.

Evidently these production functions are also not of the variety that have been used in empirical studies at the level of individual firms,[5] since all relevant firms, or at least all relevant efficient firms, are included in the observations from which these coefficient estimates are obtained.

In some ways these production functions are reminiscent of Alfred Marshall's concept of a "representative firm."[6] Here, however, the referent is rather to "representative efficient firms." Note that the

---

[1]I.e., identically the same, apart from a multiplicative constant. We are referring to the case which Sato [33], pp. 3-8, refers to as the Solow-Fisher (exact) aggregation problem. These conditions may be relaxed under a variety of assumptions. For instance, if one is willing to assume that all firms have extended Cobb-Douglas functions in the sense of [13], but not necessarily the same degree of homogeneity, then the aggregate function is also Cobb-Douglas in this same extended sense. See [14]. Condition (ii) must be satisfied in any case, since otherwise the output for given inputs will not be maximal.

[2]Cf., the discussion in Farrell [18].

[3]Again vide Farrell [18] who argues that this theoretically attainable efficiency lacks operational significance for practical applications. On the other hand, we ought to note that new constraints may be adjoined in the above programming models whenever such a priori theoretical boundaries are known to be applicable to particular firms.

[4]We are speaking generally since assumptions like continuity in the derivatives will evidently need to be modified.

[5]Cf., Johnston [24].

[6]See the discussion in Stigler [38].

plural is required insofar as there is more than a single facet. The continuum within each facet is then representative of the efficiency for which the originally observed efficient firms serve as referents.

Note, in particular, that we cannot generally average across facets to obtain a reduced set (e.g., a single set) of representative coefficients[1] without losing the property of efficiency associated with the original estimates.[2] In certain circumstances, however, we might want to regard this as being the production function associated with one "super-firm." This might be useful when, for instance, all of the underlying entities are participants in a single government program with an administrator who would like to obtain estimates of the costs which are to be expected with efficient operations.

Normally such costs would be obtained from actual or projected market data such as forecasts of teacher salaries in particular school districts or costs of supplies and so on. Here, however, we want to relate our developments to another type of duality[3] wherein $C(y,p)$ is a cost function to be determined via

$$C(y, p) = \min p^T x \text{ for } x \varepsilon L(y)$$

where

$$L(y) \equiv \{x : \text{at least the output vector } y \text{ is produced}\}.$$

---

[1]General guidelines for aggregation and disaggregation will be supplied in the next section for use when various kinds of cross-comparisons are wanted.

[2]This is also true for the interpretations of results derived from cross section data as in the interpretations of Bronfenbrenner and Douglas discussed by Sato [33] in Chapter 8 and hence is also a deficiency of the envelope and interfirm-intra firm production function interpretations of Bronfenbrenner and Reder discussed by Sato in this same Chapter.

[3]We have elsewhere suggested that this might better be associated with the branches of mathematics referred to as "transform theory." See [15].

In other words L(y) is the point to set mapping $y \rightarrow L(y)$. For instance, in Figure 1, it is obtained via the isoquant associated with y = 1 so that in this case L(y) designates all of the points $x = (x_1, x_2)$ associated with input combinations on or to the northeast of this isoquant. In any event, from a knowledge of these relations, the cost function C(y, p) is then to be obtained via the indicated minimization where p is a price vector with component p_i representing the "price" per unit x_i, the amount of the ith factor input.

The idea is to relate the above developments, including Farrell Efficiency, to another strand of theoretical work in economics[1] emanating from the original formulations by Samuelson [32] and Shephard [36].[2] The latter work, like other parts of micro-theory, has been developed with respect to decision making at the individual (firm) level. Thus, following Shephard [37] we may formulate the problem of obtaining a cost function from an already known production relation by parametrically varying p for each y to produce the portion of the production function associated with the L(y) for each y. Then parametrically varying y produces the entire cost function C(y, p) from the transformation indicated in the expressions previously displayed.

---

[1]See [22.1], [22.2] and [22.3] for detailed and extensive discussions.

[2]We have elsewhere shown how Farrell's efficiency measure can be explicitly related to Shephard's gauge function. See [15].

In our case we want our production function to be empirically based. That is, we want our production function to be based on observed input output values and estimates derived from them such that no firm from the observation set has a larger output for any inputs that may be specified. Also no non-negative combination of these firms can have a larger output when extrapolations or interpolations from the original observations must be undertaken for the specified input values.

With this definition we can then proceed to obtain the wanted cost function as follows. Let y be some prescribed output (a scalar) and let a_s, the s^{th} row of the matrix A represent the set of coefficients associated with s^{th} efficient facet estimated from the data as described in the preceding section. Let P be a matrix with its row vectors P_j representing the observational data for each of the original j = 1, ..., n firms.[1] Then on this definition of optimality (i.e., relative efficiency of the production frontiers) we can obtain the wanted cost function from the following formulation,

$$\text{(18.1)} \qquad \begin{aligned} &\min \ p^T x \\ &\text{with} \\ &\qquad A\,x \geq I\,y \\ &-P\lambda + I\,x = 0 \\ &\quad \lambda \geq 0, \end{aligned}$$

in which I is the identity matrix so that (1) $I_x = P\lambda$, $\lambda \geq 0$ assures us that we will be deriving our production function from empirically based observations, and (2) $a_s x \geq y$ together with the minimizing objective assures us that we will always be on an efficient frontier.

---

[1]One can omit the P_j which are not efficient or else one can adjust and bring them into the efficient set in the manner to be indicated in the next section.

The (mathematical programming) dual to (18.1) is

$$\max y\eta^T e$$

$$\text{with } \eta^T A + u^T I = p^T$$

(18.2) $$- u^T P \leq 0$$

$$\eta^T \geq 0 ,$$

where e is a column vector with unity for all its elements. Via the duality theorem of mathematical programming, we then have

(19.1) $$p^T x \geq y\eta^t e$$

for all x, λ and η, u which satisfy the constraints and

(19.2) $$p^T x^* = y\eta^{*T} e,$$

at an optimum. In other words,

(19.3) $$C(y, p) = y\eta^{*T} e = p^T x^*$$

is the required (minimizing) cost function, which varies with each choice of y and p.

Here we have proceeded from the production function to the cost function but, of course, we could also have proceeded via the opposite course. The latter is an "in principle" statement only, however, since, as noted in the introduction, many of the inputs and outputs in public sector applications are not easily priced or costed without recourse to arbitrary and ex cathedra procedures and assumptions. Thus we shall prefer to continue from the production rather than the cost side after observing that additional constraints may need to be adjoined to (18.1) or (18.2) to meet a variety of legal or institutional requirements.[1]

---

[1]The adjunction of such constraints can cause trouble for the Shephard-Samuelson transforms, however. See [15].

8. Efficiency and Multiple Output Production

Having exhibited how our production function and its estimation differ from others which have been customarily employed, we also ought to examine ways in which we might extend as well as utilize received constructs from micro-economics. That is, we ought to be able to use this new kind of production function to obtain new results even while attending to the need for maintaining contact (as in the preceding section) with already available theorems and results from micro-economics.

For this purpose we now turn to a new extension of the customary definition of efficiency as the ratio of output to input for the single input-output case (usually with inputs and outputs measured in the same units)[1] to the case of multiple inputs and outputs. We shall do this, moreover, in a way that leads to a generalization of the preceding developments into the case of multiple output firms in a very natural way. This multiple output case, it may be recalled, is likely to be prominent in the public policy (governmental program) areas which are of interest to us. We also want to do this in a way that releases us from the need for relying on data such as market prices which are of only limited value (and availability) for many of these applications.

To achieve these objectives we note first that if we were to weight each firm's inputs and sum, we could get a scalar measure that we might refer to as a "virtual input." Similarly, we could weight the firm's outputs and sum, to get a "virtual output." Then we could take the resulting ratio to be

---

[1] See, e.g., Van Nostrand's Scientific Encyclopedia [40].

a "virtual efficiency" for this firm, provided its value was between zero and one. Clearly there are an infinite number of possible choices of such weights and their consequent measures of efficiency.

To free our definition of efficiency from arbitrary choices of these weights and to guarantee each firm the highest efficiency rating it can receive from observational data, while rating all firms with efficiency between zero and one using these same weights, we define efficiency for a designated firm as the maximum virtual efficiency obtainable with non-negative weights that impute an efficiency between zero and one to every firm in the comparison set. In formulas, we have

$$\max h_o = \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}}$$

(20) subject to

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1, \quad j=1, \ldots, n$$

where $u_r, v_i \geq 0$ are weights applied simultaneously to the $j = 1, \ldots, n$ firms observed to utilize common inputs $i = 1, \ldots, m$ in amounts x_{ij} and produce common outputs in amounts y_{rj}, $r=1, \ldots, s$. The subscript zero is reserved for the designated firm.

To recapitulate, each firm will have its efficiency rated relative to every other firm producing the same outputs from the same kinds of inputs. This is to say (again) that the efficiency rating will be relative to the comparison set. With this understanding the resulting rating is secured from the indicated ratio by means of an optimal (maximizing) choice of non-negative weights in such a way that no firm is rated more than 100% efficient.

In our case all input and output observations are positive and so it is not necessary to deal with indeterminate or nonsolution possibilities. In any case the above formulation is an extended (nonlinear) formulation of an ordinary fractional programming problem for which we have elsewhere supplied a complete theory.[1] This means that we can deal with such possibilities in already known ways and hence need not turn aside to deal with issues such as possible indeterminacies in any case.

---

[1]See [10] and [7].

Thus we now proceed to relate the above measure to other parts of our analysis by utilizing obvious manipulations to replace it with the following formulation for an "inefficiency measure":

$$\min f_o = \frac{\sum_{i=1}^{m} v_i x_{io}}{\sum_{r=1}^{s} u_r y_{ro}}$$

(21) subject to

$$\frac{\sum_{i=1}^{m} v_i x_{ij}}{\sum_{r=1}^{s} u_r y_{rj}} \geq 1, \quad j=1, \ldots, n$$

$$v_i, u_r \geq 0.$$

Now we propose to replace this formulation with an <u>ordinary</u> linear programming problem as follows. First consider

$$\max z_o$$

subject to

(22)

$$-\sum_{j=1}^{n} y_{rj}\lambda_j + y_{ro} z_o \leq 0, \quad r=1, \ldots, s$$

$$\sum_{j=1}^{n} x_{ij}\lambda_j \leq x_{io}, \quad i=1, \ldots, m$$

$$\lambda_j \geq 0, \quad j=1, \ldots, n .$$

Because (22) is an ordinary linear programming problem it has a linear programming dual which we can write as follows:

$$\text{Min } g_o = \sum_{i=1}^{m} \omega_i x_{io}$$

subject to

(23)
$$-\sum_{r=1}^{s} \mu_r y_{rj} + \sum_{i=1}^{m} \omega_i x_{ij} \geq 0$$

$$\sum_{r=1}^{s} \mu_r y_{ro} = 1$$

$$\mu_r, \omega_i \geq 0 .$$

Because of the condition $\sum_{r=1}^{s} \mu_r y_{ro} = 1$ in (23) one can recognize that it is equivalent to an ordinary linear fractional programming problem. In fact, utilizing the theory of linear fractional programming with the transformation

(24)
$$\omega_i = tv_i, \; i=1, \ldots, m$$

$$\mu_r = tu_r, \; r=1, \ldots, s,$$

which, with $t > 0$, gives explicitly

$$\min f_o = \frac{\sum_{i=1}^{s} v_i x_{io}}{\sum_{r=1}^{s} u_r y_{ro}}$$

(25) subject to

$$\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} \geq 0, \quad j=1, \ldots, n$$

$$v_i, u_r \geq 0 ,$$

as the equivalent of (23). By very evident manipulations, however, we can see that (25) is the same as (21). Q.E.D.

We are now in an advantageous position from several standpoints. We have a completely symmetric definition of efficiency which generalizes single output ratio definitions not only in economics but in engineering and other natural sciences.[1] We do not need to solve the nonlinear (and nonconvex) problems in which these definitions are formalized. We need only solve the ordinary linear programming problem (23) in order to obtain both the optimal f_o^* or h_o^* and the weights $v_i^*, u_r^* \geq 0$.

This can be seen by observing that multiplication of all expression in (25) by any $t > 0$ leaves the optimal value of f_o unaffected. Thus, choosing $t^* = 1$ we have at once

$$f_o^* = g_o^* = z_o^* \qquad (26.1)$$

and therefore

$$h_o^* = 1/z_o^* = \text{TEF}(P_o) \qquad (26.2)$$

as well as the wanted (relative) weights.

Thus nothing more is required than the solution of (23) or (22) in order to determine whether $f_o^* > 1$ or, correspondingly, whether $h_o^* < 1$, with efficiency prevailing if and only if

$$f_o^* = h_o^* = 1. \qquad (26.3)$$

In this way we have dispensed with condition ii in (13), as promised. Hence, our use of the single characterization for $\text{TEF}(P_o) = h_o^* = 1/z_o^*$ is justified when only this measure efficiency is wanted.

---

[1] See, e.g., [40].

For production function and other analyses, however, we may need to retu to condition (ii) in (13) for various adjustments. To indicate some of these possibilities, we should first recognize that (22) is a generalization of (14.1) in which (7.1) is replaced by new observational vectors

$$(27) \qquad P_j = \begin{pmatrix} Y_j \\ X_j \end{pmatrix}, \quad j=1, \ldots, n,$$

wherein the subvector Y_j contains the observed output values y_{rj}, $r=1, \ldots, s$ for its components and the subvector X_j contains the observed input values x_{ij}, $i=1, \ldots, m$.

Before proceeding further we need to make the following remarks. Because of the presence of multiple outputs we have omitted the use of norms and thus record these y_{rj} and x_{ij} values directly.[1] I.e., we do not normalize on output as we did in (7.1) and (7.2). In fact, the concept of an isoquant becomes ambiguous under such multiple output analysis and indeed the concept of a production function gives way to more general concepts such as "production possibility sets" and the "activities" to which they are related.[2] Efficiency and related concepts continue to maintain, however, as do our methods of securing estimates of the activity coefficients from empirical data via the mathematical programming dual (22).[3]

Adjustments are needed in both concept and method for these more complex multiple output situations, of course, but these are minor

---

[1]The norming occurs in a more recondite form in the final constraints of (23).

[2]Cf., e.g., Arrow and Hahn [4], p. 52 ff.

[3]I.e., (23). We note that there is also a duality theory for fractional programming which remains to be exploited for (20) and (21). See Bector [5], Jaganathan [23], and Schaible [34].

and relatively straightforward. Hence, we will not deal with them here in order to enable us to devote our remaining attention to the kinds of adjustments that are needed when we want to ensure that the resulting (adjusted) observations are on the relevant efficiency frontiers.

Consider therefore the following reformulation of (22):

$$\text{Max } z_o$$

with

$$(28) \qquad \sum_{j=1}^{n} Y_j\lambda_j + Y_o z_o \leq 0$$

$$\sum_{j=1}^{n} X_j\lambda_j \leq X_o$$

$$\lambda_j \geq 0, \quad j=1, \ldots, n .$$

Let its optimal solution be represented

$$(29) \qquad z_o^*, \; s^{*+}, \; s^{*-}, \; \lambda_j^*, \quad j=1, \ldots, n ,$$

where s^{*+} represents a vector of non-negative slack associated with the output additions and s^{*-} represents a vector of non-negative slack associated with input subtractions. If $z_o^* > 1$ or if any component of s^{*+} or s^{*-} is positive then via (13) the efficient frontier of the production possibility surface has not been attained.

If the observations for this firm are to be utilized in further analysis of a micro-theoretic variety it is necessary to bring them onto the surface in the following manner. All output must be scaled upward by the multiple $z_o^* \geq 1$. Also all slack must be reduced to zero by increasing the outputs or decreasing the inputs for which non-zero slack appeared in (29).

To make these statements more precise and to show that a point on the efficiency frontier is thereby produced, we utilize the data of (29) and replace (28) with the following new problem:

$$\text{Max } \hat{z}_o$$

with

$$(30) \qquad \begin{aligned} -\sum_{j=1}^{n} Y_j\hat{\lambda}_j + (Y_o z_o^* + s^{*+})\hat{z}_o &\le 0 \\ \sum_{j=1}^{n} X_j\hat{\lambda}_j &\le X_o - s^{*-} \\ \hat{\lambda}_j &\ge 0 \,. \end{aligned}$$

We shall refer to (30) as the "varied problem" and show that the thus adjusted observations satisfy the conditions for efficiency - (13) ff - as follows. Evidently we can secure $\hat{z}_o^* = 1$ since $\hat{z}_o^* = 1$ together with (29) gives us the already secured optimal solution to (28). Now suppose we could have $\hat{z}_o^* > 1$ in (30). This would yield

$$\begin{aligned} -\sum_{j=1}^{n} Y_j\hat{\lambda}_j^* + Y_o\hat{z}_o^* z_o^* &\le -\sum_{j=1} Y_j\hat{\lambda}_j^* + (Y_o z_o^* + s^{*+})\, \hat{z}_o^* \le 0 \\ \sum_{j=1}^{n} X_j\hat{\lambda}_j^* &\le X_o - s^{*-} \le X_o, \end{aligned}$$

since s^{*+} and s^{*-} are both non-negative. Evidently, the expressions on the left then satisfy the "unvaried problem" (28) with $z_o^* \hat{z}_o^*$ in place of z_o^* and $\hat{\lambda}_j^*$ in place of λ_j^*. However, then also

$$\text{Max } z_o \ge z_o^* \hat{z}_o^* > z_o^*$$

when $\hat{z}_o^* > 1$. But $z_o^* = \text{Max } z_o$, by hypothesis. Thus a contradiction occurs which can be resolved only by assuming $\hat{z}_o^* = 1$ as the optimal value for the varied problem (30).

Now we want to show that the optimal solution, λ_j^*, j=1, ..., n to the unvaried problem (28) is an optimal solution to the varied problem (30) with zero slack, i.e., the vectors $\hat{s}^{*+}$ and $\hat{s}^{*-}$ have zeros in all components as required for efficiency. First, we show via (29) and (30) that

$$-\sum_{j=1}^{n} Y_j\lambda_j^* + Y_o z_o^* + s^{*+} = 0$$

$$\sum_{j=1}^{n} X_j\lambda_j^* \qquad\qquad = X_o - s^{*-}$$

is a feasible solution of the varied problem with $\hat{z}_o = 1$. That is

$$-\sum_{j=1}^{n} Y_j\lambda_j^* + (Y_o z_o^* + s^{*+})\ \hat{z}_o = 0$$

$$\sum_{j=1}^{n} X_j\lambda_j^* \qquad\qquad = X_o - s^{*-}$$

with $\hat{z}_o = 1$. However, with $\hat{z}_o = 1$ the λ_j^* from the original unvaried problem continue to satisfy the constraints and are optimal for the varied problem, too, after the z_o^*, s^{*+} and s^{*-} adjustments to efficiency are effected. Q.E.D. In short, the indicated adjustments do, in fact, always bring the original observations to the relevant efficiency frontier. No new computations are required after the z_o^*, s^{*+} and s^{*-} adjustments are effected for the original Y_o, X_o data for the efficiency comparisons we may subsequently want to make since all observations are on the efficiency surface as required.

9. Concluding Remarks

The adjustment procedures we have just described can be put to use in a variety of ways. They may be used, for instance, to distinguish between managerial decisions and other sources of efficiency as in the Data Envelopment Analyses which were discussed in our introduction. See [30]. They can be used in other ways as well. For instance Farrell's technical efficiency concept was employed by Carlsson [6] in an attempt to measure the (relative) efficiency of Swedish firms by reference to Leibenstein's concept of "X-efficiency."[1] The developments set forth in our paper should be helpful in meeting criticisms of such work as was leveled at Carlsson by Førsund and Hjalmarson in [21]. This is by virtue of the fact that ou formulations can handle large numbers of constraints and almost limitless numbers of observational vectors. Thus in this way we can go a considerable distance in admitting different capital vintages, say, or different classes of labor for explicit treatment as possible sources of variations in efficiency.[2]

Data availability at the level of individual decision-making units may cause problems, although this is likely to be more serious in dealing with private enterprise statistics than the kind of public programs which form the main concern of the present paper. For cases in which aggregations or disaggregations are wanted, however, as for the synthesis of a variety of different production functions from the same underlying data, we can provide the following general guidelines.

---

[1]See [27] and [38]. Although this view has recently come under severe criticisms by Stigler [38], no empirical evidence was submitted in Stigler's argument to refute Leibenstein's claim of the existence of such inefficiency.

[2]Actually it was Farrell's criticism of Colin Clark's omission of some factors of production (to account for some of Clark's results) that appeared to motivate Farrell's original work in [18].

In synthesizing from among the many possibilities for defining such aggregate (or subaggregate) functions, properties like the following need to be kept in mind. A "representative" efficient surface (e.g., the unit isoquant in the single output case) consists of facets, each of which is a simplex. Each such simplex spans a convex cone. Thus when considering the aggregation, i.e., the summing, of two or more inputs, the related efficient aggregate output will be the sum (or aggregate) of the efficient outputs of the individual firms as long as they are in the same cone. If the individual inputs are from different cones, the aggregate input may lie in any one of the individual cones and it can lie in other cones as well. In this case (i.e., at least two inputs from different cones), the efficient aggregate output will exceed the sum of the efficient individual outputs by virtue of the strict sub-additivity[1] of our efficient production function for input vectors from different cones.

Other relations of disaggregation and aggregation may also be employed, of course, according to purposes that might be served thereby. Inter alia, none of the above is intended to imply that customary approaches to the study of empirical production relations must necessarily be abandoned. It means rather that we now have numerous other alternatives that can be shaped and applied for different purposes as the contexts may suggest.

---

[1]See [9], p. 284 ff.

Some of the advantages and possibilities for use having been indicated at various points in the text, we can best conclude with some of the possible limitations. In addition to potential lack of data availability at individual firm levels, we have also noted that our approach has been mathematical (i.e., deterministic) rather than statistical in character. We have already indicated some of our reasons for delaying a treatment of this subject since we propose to come at it in an entirely different way via the techniques that we refer to as Data Envelopment Analysis. For the present, however, we must admit that our treatment of these statistical considerations has been weak since, in principle at least, one must admit to the presence of statistical error even when all-inclusive data are at hand. Of course other alternatives to Data Envelopment Analysis may also be explored as in the treatments by Timmer [39] among others, who, following up a suggestion by Aigner and Chu [3], has employed a "chance constrained programming" formulation -- but without really exploiting anything like the full range of formal theory that is available from that quarter.[1]

Finally we turn to the measure of efficiency. We have repeatedly noted its restriction to relative efficiency. For the kinds of public policy applications with which we are concerned, this provides an approach to controlling and evaluating managerial behavior. By this we mean that managers who fail to meet the indicated efficiency level can at least be confronted with the evidence and required to justify their departure from what "the evidence" suggests is attainable. In this manner a basis for further learning can also be provided since any special or ameliorating circumstances revealed in the course of such inquiries can then be introduced as constraints into future models and measurement results.

[1] Cf., e.g., [8].

As movement between industries or even regions becomes a pertinent issue, however, the case for this measure of efficiency begins to weaken. In golfing terminology it is, so to speak, a measure of "distance" rather than "direction" with respect to what has been (and might be) accomplished. As such it is inferior to "profit" which measures direction as well as distance and possibly other aspects of attainment, including, e.g., the best among all available alternatives, whenever it is really applicable. On the other hand, most public endeavors (schools, police, etc.) are not free to transfer resources simply because some other more attractive alternatives exist. Hence the fact that their resources are committed in an indicated direction suggests that such measures of accomplishment as we have considered here are appropriate and they are certainly needed for at least some purposes of policy evaluation and prescription.

APPENDIX

In all of our treatments in this paper an important consideration is the convexity of the set of production possibilities. There may, of course, be many different measures for the amount of inputs or outputs and so we need to consider how the resulting measure of efficiency may be altered as these input and output measures are altered.

For clarity we consider the single output case. There may even be transformations of the n+1 coordinates of inputs plus outputs into other coordinate systems (with n+1 coordinates) in which, say, a nonconvex set of production possibilities is transformed into a convex set. See section 1 of the paper. Provided that such a transformation is "bicontinuous," [1] however, boundary points remain boundary points and interior points remain interior points. Thus when there are two different coordinate representations both of which have convex sets of production possibilities, efficient points remain as such in both systems and the same is true for nonefficient points. Hence the characterization of any point as efficient or nonefficient will remain invariant under any such transformation. The numerical magnitude may, of course, change. Specific examples can be given, but it is important to note that even this numerical value remains invariant under any "homothetic" transformation. The latter, which refers to constant changes of scale, is the one that will usually be of interest [2] and hence, in this respect, too, we can build on the important opening provided by Farrell [18] for empirical work and related policy applications. In fact, as we elsewhere show [30] new techniques of data analysis can then be developed for added use in cases where more customary statistical and econometric techniques may otherwise fail to provide what is wanted.

---

[1] See, e.g., J. L. Kelley [25].

[2] See the discussion in Shephard [37].

Bibliography

[1] Afriat, S.N., "Efficiency Estimation of Production Functions," International Economic Review, 13, No. 3, Oct. 1972, pp. 568-583.

[2] Aigner, D.J., T. Amemiya, P.J. Poirier, "On the Estimation of Production Frontiers: Maximum Likelihood Estimation of the Parameters of a Discontinuous Density Function," International Economic Review, XVII, No. 2, June 1976, pp. 377-396.

[3] Aigner, D.J. and S.F. Chu, "On Estimating the Industry Production Function," American Economic Review, LVIII, Sept. 1968, pp. 826-839.

[4] Arrow, K.J. and F.H. Hahn, General Competitive Analysis, (San Francisco: Holden-Day, Inc. 1971).

[5] Bector, C.R., "Duality in Linear Fractional Programming," Utilities Mathematica, 4, 1973, pp. 155-168.

[6] Carlson, B., "The Measurement of Efficiency in Production: An Application to Swedish Manufacturing Industries, 1968" Swedish Journal of Economics, 1972.

[7] Charnes, A. and W.W. Cooper, "An Explicit General Solution in Linear Fractional Programming," Naval Research Logistics Quarterly, 20, No. 3, Sept. 1973.

[8] Charnes, A. and W.W. Cooper, "Chance Constrained Programming," Management Science, Oct. 1959.

[9] Charnes, A. and W.W. Cooper, Management Models and Industrial Applications of Linear Programming, Vol. I, (New York: John Wiley and Sons, Inc. 1961).

[10] Charnes, A. and W.W. Cooper, "Programming With Linear Fractional Functionals," Naval Research Logistics Quarterly, 9, Nos. 3 and 4, Sept.-Dec. 1962, pp. 181-185.

[11] Charnes, A., W.W. Cooper and K.O. Kortanek, "Semi Infinite Programming Differentiability and Geometric Programming Part I: With Example Applications in Economics and Management Science," Journal of Mathematical Sciences, R.S. Varma Memorial Volume, 6, 1971, pp. 19-40.

[12] Charnes, A., W.W. Cooper and E.L. Rhodes, "Expositions, Interpretations and Extensions of Farrell Efficiency Measures," Research Report (Pittsburgh: Carnegie-Mellon University School of Urban and Public Affairs, April 30, 1975.)

[13] Charnes, A., W.W. Cooper and A. Schinnar, "A Theorem on Homogeneous Functions and Extended Cobb-Douglas Forms," Proceedings of the National Academy of Sciences, Oct. 1976, Vol. 73, No. 10.

[14] Charnes, A., W.W. Cooper and A.P. Schinnar, "An Extended Cobb-Douglas Form for Use in Production Economics." Pittsburgh, Pa. Carnegie-Mellon University, School of Urban and Public Affairs, Institute of Physical Planning, Research Report (mimeo). Dec. 1975.

[15] Charnes, A., W.W. Cooper and A.P. Schinnar, "Transforms and Approximations in Cost and Production Function Relations," Research Report CCS 284 (Austin, Texas: University of Texas Center for Cybernetic Studies, January 1977).

[16] Dantzig, G.B., Linear Inequalities and Related Systems (Princeton: Princeton University Press, 1963).

[17] Duffin, R.J., E.L. Peterson and C.M. Zener, Geometric Programming (New York: John Wiley & Sons, Inc., 1967).

[18] Farrell, M.J., "The Measurement of Productive Efficiency," Journal of the Royal Statistical Society, Series A, Part III, 1957, pp. 253-290.

[19] Farrell, M.J. and M. Fieldhouse, "Estimating Efficient Production Functions Under Increasing Returns to Scale," Journal of the Royal Statistical Society, Series A, Part II, 1962, pp. 252-267.

[20] Førsund, F.R., Studies in the Neo-Classical Theory of Production (Oslo: University of Oslo Institute of Economics Memorandum, 1974).

[21] Førsund, F.R. and L. Hjalmarsson, "Comment on Bo Carlsson's 'The Measurement of Efficiency in Production: An Application to Swedish Manufacturing Industries, 1968,'" Swedish Journal of Economics, 1974.

[22] Intrilligator, M.D. and D.A. Kendrick, eds., Frontiers of Quantitative Economics, Vol. II. (Amsterdam: North Holland Publishing Co., 1974).

[23] Jaganathan, R., "Duality for Nonlinear Fractional Programs," Zeitschrift für Operations Research, Band 17, Seite 1-3, Würzburg: Physica-Verlag, 1973.

[24] Johnston, J., Statistical Cost Analysis (New York: McGraw-Hill Book Co., Inc., 1960).

[25] Kelley, J.L., General Topology (New York: Van Nostrand, 1955).

[26] Kuhn, T.S., The Structure of Scientific Revolutions, 2nd ed. International Encyclopedia of Unified Science 2, No. 2 (Chicago: The University of Chicago Press, 1970).

[27] Leibenstein, H., "Allocative Efficiency vs. X-Efficiency," The American Economic Review, June 1966, pp. 392-415.

[28] Leibenstein, H., Beyond Economic Man (Cambridge: Harvard University Press, 1976).

[29] Lemke, C.E., "The Dual Method of Solving the Linear Programming Problem," Naval Research Logistics Quarterly, Vol. I, 1954, pp. 36-47.

[30] Rhodes, E., Data Envelopment Analysis and Related Approaches for Measuring the Efficiency of Decision Making Units. (Pittsburgh: Carnegie-Mellon University, School of Urban and Public Affairs, in process, 1977 completion scheduled).

[31] Robinson, A., Non-Standard Analysis, Studies in Logic and the Foundations of Mathematics (Amsterdam: North Holland Press, 1966).

[32] Samuelson, P.A., "Prices of Factors and Goods in General Equilibrium," The Review of Economic Studies, 21, 1953-1954, pp. 1-20.

[33] Sato, K., Production Functions and Aggregation (Amsterdam: North Holland Publishing Co., 1975).

Bibliography (Cont.)

[34] Schaible, S., "Parameter-free Convex Equivalent and Dual Programs of Fractional Programming Problems," Zeitschrift für Operations Research, Band 18, Seite 187-196, Würzburg: Physica-Verlag, 1974.

[35] Seitz, W.D., "Productive Efficiency in the Steam-Electric Generating Industry," Journal of Political Economy, Vol. 79, No. 4, 1971, pp. 878-886.

[36] Shephard, R.W., Cost and Production Functions (Princeton: Princeton University Press, 1953).

[37] Shephard, R.W., The Theory of Cost and Production Functions (Princeton: Princeton University Press, 1970).

[38] Stigler, G.J., "The Xistence of X-Efficiency," The American Economic Review, March 1976, pp. 213-216.

[39] Timmer, C.P., "Using a Probabilistic Frontier Production Function to Measure Technical Efficiency," Journal of Political Economy 79 (4), 1971, pp. 776-794.

[40] Van Nostrand's Scientific Encyclopedia, 5th ed., Douglas M. Considine Ed. in Chief (New York: Van Nostrand-Rineholt, 1976).