«ΣΠΟΥΔΑΙ», Τόμος 45, Τεύχος 1ο-2ο, Πανεπιστήμιο Πειραιώς / «SPOUDAI», Vol. 45, No 1-2, University of Piraeus

THE EVOLUTION OF NON-PARAMETRIC FRONTIER ANALYSIS METHODS: A REVIEW AND RECENT DEVELOPMENTS

By

Antreas D. Athanassopoulos
University of Warwick

Abstract

The assessment of performance of activity units such as bank branches, retail outlets, sales forces in the profit-making sector and schools, local authorities and hospitals in the non-profit sector are given increased attention by the econometric and operational research communities in recent years. A relatively new methodology has been developed based on the use of linear programming for assessing performance of activity centres. This paper provides a state of the art review concerning the basic models and also recent trends of this methodology. (JEL C44, D24)

1. Introduction

During the last two decades enormous attention has been given to the assesment and improvement of performance of production systems. Economic activities at the firm, industry, region or nation level are affected by the world-wide trend for improved performance. National economies for example, Japan, have gained economic advantage during these decades due to their ability to improve performance in their manufacturing and service delivery systems.

On the other hand, the continuous economic recession in the western world, the failure of the welfare state of the seventies in Europe and the subsequent failures of the liberal and neoliberal systems of the eighties to control public spending and public deficit put enormous pressures on profit and not-for-profit organisations for improving performance as a means to long run viability[1]. More recently the collapse of the socioeconomic structure in the excommunist countries brought up the question of performance in a previously unknown scale. The question of improving performance has gained popularity among various political parties and now various initiatives can be found discussing the issue of performance as a distinct political, economic and social concept.

In the UK, particularly, during the 1980s there erupted a concern for *accountable management* within public sector organisations. Since then a new generation of professionals and academics has flourished with particular focus on the assessment of productive efficiency of systems. On the other hand, accountability and performance measurement in the private sector has received increased attention spurred in part by unseccessful experiences of auditing bodies (accounting firms) to uncover the true performance prospects of many profit making organisations, (e.g. see the Poly Peck and BCCI fraud cases in the UK).

The revival of the performance measurement culture mainly in the public but also in the private sector has brought closer previously unconnected disciplines that are by nature involved with the assessment of performance. Clearly, the assessment of performance has *political, economic, accounting and management science* dimensions which could be integrated to improve the way performance is assessed.

The rest of the paper is organised as follows. A review of performance measurement is made emphasising its multi-dimensional nature. It is argued that individual disciplines can address the question of assessing performance in part and thus a framework needs to be developed for integrating the strong features of different disciplines into a common performance measurement discipline proposed. This framework is called *frontier analysis* and includes in its development stochastic and deterministic variants. The paper concentrates on the nonparametric and deterministic aspects of *frontier analysis* with particular emphasis on its technical aspects and recent developments in this area.

2. A Tour in the History of Performance Measurement

For some authors the history of analysing performance of organisations is dated back to Plato's and Aristotle's discussions about the effectiveness of different military organisations, Hoagland (1964). Leonardo da Vinci in the fifteenth-century also studied performance questions concerning labour effort in shovelling. The previous two references are intended to show that the concept of performance is an old problem in the history of sciences and philosophy. However, the glory belongs almost exclusively in the post nineteenth century and in particular to F. Taylor who has been characterised as the father of scientific management[2]. Aside to the controversy about its originator, scientific management represents an attempt at improving the efficiency of various operating systems using laws and methods from the natural sciences.

Since the development of scientific management other related disciplines have advanced towards defining and considering the concept of performance in management from their own perspective. Next, the political, accounting, economic and management science views on performance measurement are discussed.

2.1. Politics and Performance

The assesment of performance of systems has inherent political dimension as it reflects the purpose and mission of the system. Political institutions seek to improve and enhance the performance of societies and economies based on a set of ideological principles. For many political scientists assessment of performance of institutions should emphasise primarily issues related with *freedom, access to power, decision making and rationalisation.* In this conception economic achievements in societies without democratic freedoms are not considered of any real value as it is argued that in the long run the performance of the system will decline and eventually collapse. For example, the economic performance of excommunist countries was slowed down by the lack of real political democracy.

Another area of linkage between politics and performance emanates from the concept of *political choice.* This can be demonstrated using the privatisation of public utilities in the UK in the eighties. Advocates of privatisation argue that the public control of telecommunications, electricity, water, etc. prevent them from operating efficiently and therefore should be run under private control. Competition and market conditions are expected to stimulate the economy, reduce costs, thereby, benefiting the tax payer who will not have to contribute through taxation to the potentially inefficient operation of companies like British Telecom. Performance can be used, therefore, to support decisions of political nature.

The opponents of privatisation, however, also use performance related arguments to object to privatisation. For example, Mayston (1993) argues that the political decision to sell public assets has short run benefits on public finance but in the long run the public "loses" the opportunity to gain financially from the very high profits of these companies (e.g. £96 per second profits for British Telecom).

2.2. Accounting and Performance

The concept of organisational accountability in both profit and not-for-profit organisations constitutes another dimension of performance. Accountability has strong political origins as it is the process that informs shareholders[3] on the propriety of decisions made in organisations. Historically, the accounting profession has been employed to generate information related to organisational performance.

Booth and Cocks (1990) state that the accounting profession has traditionally been viewed as a neutral purveyor of the facts. Accounting's role in any type of institutions is growing over time. Critiques and advocates give various explanation for this phenomenon. The most important issue that arises from this debate, however, is whether accounting information provides sufficient evidence on the performance status of organisations. Evidence obtained from the accounting literature emphasise that the current accounting practices give little assurance to shareholders on whether or not companies are performing adequately.

An empirical study in the private sector by Citron and Taffler (1992) found no correlation between whether or not an audited firm received a going concern qualification and whether or not it subsequently failed in the next 12 months. The general case of creative accounting is well discussed in the accounting literature as an ongoing and growing problem of the accounting profession, Griffiths (1992). Similar messages can also be found in the use of accounting practices in the public sector. For example, the demanding data requirements for implementing the Resource Management Initiatives (RMI) of the national health system in the UK have very high cost implications for developing appropriate information systems without substantial performance returns, (National Audit Office, 1992).

2.3. Economics and Performance

The economic approach to efficiency is perhaps the most elegant one. The reason for that relates to the axiomatic definition of economic phenomena, such as production, and the subsequent examination on whether these theoretical models are supported by real life facts. Efficiency is perceived in economics as the outcome of comparing the actual output of productive units againts a theoretical defined maximum output given the resources used. At the theoretical level this is represented by the notion of the *production function* which in short represents an extreme relationship between inputs and outputs and accounts for the maximum obtainable amount of output for a given level of input and *vice versa.*

The closest association between efficiency and economics can be found in the theory of production. Looking back in history one can find a number of key contributors that affected in one way or another the development of what we shall introduce later as *frontier analysis.* Chambers (1990) in his monograph on production theory uses the agricultural experiments (1820-1830) of Von Thuenen as a starting point for production theory whilst recognising Moore (1929) as one of the originators in using statistics to examine economic phenomena such as the marginal productivity theory.

According to Lovell (1993) a departure point of efficiency studies is Knight (1933) who defined efficiency as the ratio between outputs and inputs and furthermore discussed issues related with the selection of inputs/outputs for assessing efficiency. Chambers (1990) argues that despite the earlier studies of production relationships it was only after the seminal work by Cobb and Douglas (1928) that the estimation of the production function became common place in economics.

Lovell (1993) discusses in more detail issues and problems related with the econometric methodology for assessing efficiency. One can argue that the estimation of econometric based frontiers, despite its advances during the last two decades, has yet to address the problem of selecting appropriate functional forms, the distributional problems of the inefficiency terms and the accommodation of multiple input-output cases. Schmidt (1985) and Thanassoulis (1993) discuss in more detail the pros and cons of econometric frontier estimation.

2.4. Management Science and Performance

As mentioned earlier, the development of scientific management sought to borrow from the natural sciences for improving performance of socio-economic systems. In the post second world war period scientific management was enhanced by "operational research" techniques. A large number of problems concerning resource allocation, location analysis, transportation planning, educational and health care planning and delivery were supported using tools like linear programming, project management (PERT, GANTT), decision trees, simulation, and queuing theory.

The main emphasis of operational research methods was to provide decision support for planning. These efforts, however, did not consider the possibility of using operational research techniques in a control mission for assessing organisational performance. As Charnes and Cooper (1978) note:

> "Almost no attention has been devoted to improved procedures of accountability and/or other approaches to the control of management behaviour"

In summary, performance measurement has very important political, accounting, economic and management science affiliations. The definition and assessment of performance measurement can vary from being an abstract political concept to a set of performance indices reported by accounting auditors. There is a fundamental agreement that performance measurement needs to have a quantitative component where performance is assessed by some type of ordinal or nominal scale.

3. The Evolution of Frontier Analysis

The previous discussion focused on the political, accounting, economic and management science dimension of performance measurement. As none of these disciplines can capture performance measurement in full some synthesis towards a unified framework is necessary. An attempt towards this direction is made via frontier analysis. Aigner, Lovell and Schmidt (1977) note characteristically:

> The theoretical definition of a production function expressing the maximum amount of output obtainable from given input bundles with fixed technology has been accepted for many decades. And for almost as long, econometricians have been estimating average production functions.
>
> [Aigner, Lovell and Schmidt (1977, ρ 21)]

As mentioned earlier, traditional economic approaches use theoretically justified production function and test their behaviour on real data. Data sets that do not support the prespecified production functions have two possible interpretations. Either the specified production function was inappropriate or the productive units in the analysis were very inefficient and therefore could not give a sufficient statistical fit. *Unfortunately, these types of problems do not seem to have any obvious answer.*

Farrell (1957) was the first to put forward an alternative framework for assessing productive efficiency by reverting the order efficiency assessment was pursued by traditional economics. Farrell suggested that productive efficiency should be assessed using empirical observations avoiding *a-priori* specification of functional forms. A pictorial representation of the work promoted by Farrell (1957) is given in Figure 1.

Figure 1, illustrates an example where Decision Making Units (DMUs) require two inputs for producing one output. The input quantities have been standardised per unit of output produced and therefore the example has adopted a constant[4] returns to scale assumption. Suppose that the *efficient production function* is known and given by the curve SS'. In other words, the output that a perfectly efficient firm could obtain from any combination of inputs. Let us also assume that the prices pi, p2 for the two input quantities are also known; the line AA' (pixi + P2X2 = C) has a slope equal to the ratio of the prices of the two inputs. Where C is the cost of securing one unit of output.

Let us compare DMUs P and Q. They are both on the same ray from the origin which implies that they employ the same input mix (proportions). However, unit Q produces the same output as P using only a fraction OQ/OP of the inputs used by Q. *We shall define, therefore, the ratio OQ/OP as the technical efficiency of unit Q.*

Of equal importance is to find out the extent to which a firm uses the various factors of production in the best proportions, in the light of their prices. Comparing points Q and Q' on the theoretical production function it is obvious that Q' uses the least cost input combination for producing a unit of output. The costs of production at Q' are a fraction OR/OQ of those at Q. *This ratio is defined as the price or allocative efficiency of Q. It also represents the price efficiency of all technically inefficient DMUs such as P that have been projected at point Q.*

Overall, if unit P was technically and price efficient its costs would be a fraction OR/OP of its present levels. This ratio is called the overall efficiency of unit P and can be decomposed into its technical and price efficiency components as follows: OR/OP = OQ/OP ^{x} OR/OQ.

FarrelFs work was innovative for a number of reasons:

- The need for specifying the functional form of production functions prior to estimating the productive efficiency on empirical data was relaxed,

- Efficiency was decomposed into *technical, allocative and overall* components. Later he also added a *scale efficiency* component,
- Linear programming in a performance measurement mode was used,
- The existence of multi-input and multi-output production functions was recognised without, however, providing a way of estimating them.

FarrelPs work did not find an immediate widespread use and it was Aigner and Chu (1968) and Forsund and Hjalmarsson (1979) that launched the first attempt for assessing efficiency using Farrelf s rationale.

Aside of these developments a parallel stream of economic thought was developed by Leibenstein postulating the existence of nonallocative inefficiency in production (i.e. nonoptimal mix of inputs). Frantz (1992) argues that until that time economists thought mainly about allocative inefficiency and assumed that firms were always maximising their technical efficiency due to the market's pressure. Lovell (1993) argues that there is scope for linking the literature of X-efficiency with the performance measurement literature as has evolved from the post Farrell (1957) period. Leibenstein and Maital (1992) seem to agree with this as they appreciate the potential similarities between X-efficiency and frontier analysis.

The turning point, after Farrell (1957), in the assessment of performance at the firm level came via two parallel attempts from the operational research [Charnes *et al.* (1978)] literature and the economic [Fare and Lovell (1978)] literature. A new "technique" called *data envelopment analysis* (DEA) emerged from these initial attempts opening a very wide research area which since then gained widespread development. Since then there has been a considerable expansion of the method in both theoretical and applied terms, [Vassiloglou and Giokas (1991), Giokas (1992), Athanassopoulos and Thanassoulis (1994)]. Technical reviews of the method can be found in [Boussofiane *et al.* (1991), Seiford and Lewin (1990) and Ali and Seiford (1993)].

3.1. Frontier Analysis Components

This section focuses on the development of a framework for assessing the technical efficiency of DMUs using the principles of frontier analysis. Frontier analysis is perceived as having three interrelated components.

- A systems' component
- A mathematical programming component
- A decision support component

Frontier analysis seeks to investigate the performance of productive systems which employ input factors to deliver outcomes as represented diagrammatically in Figure 2. The very nature of performance measurement is heavily influenced by the inputs/outputs identified in a production process. For example, assessing the performance of schools using as inputs the resources available at a school (no. of teachers, facilities and expenditure) and as an output the examination achievements of pupils one can assess the rate schools utilise their resources by achieving high examination results. If, however, the input list included information on the entry standards, as well as sociodemographic background of pupils one would argue that the assessment yields information concerning the value added at schools, Thanassoulis and Dunstan (1993). Apart from the nature of the inputs/outputs used for assessing performance questions concerning the appropriate number of inputs/outputs for describing an activity process can also be raised.

In the economic literature one can find extreme opinions about the role of inpu-output systems in assessing performance. For instance Knight (1933) argued that if all inputs and outputs are included in assessing the efficiency of DMUs then they will all get an efficiency of unity (100%). Knight, therefore, made a suggestion for redefining productivity using only the "useful" inputs and outputs. More recently Ray (1988) argued that the measured inefficiency may reflect the failure to incorporate all the right variables and constraints and to specify the right economic objective, of the production unit.

3.2. Production Possibilities and Efficient Frontiers

The discussion of any production activity in economic theory must draw on the notion of the *production possibility set*. The production possibility set, Φ, is in theory an unknown and therefore it will either be defined in abstract or it will be defined using observed production units. Let us suppose that we have data on a set of $j = 1, \ldots, n$ DMUs and each DMU use inputs $X \in R_+^s$ to produce outputs $Y \in R_+^m$. Therefore unit j uses amount x_{ij} of input i to produce amount y_{rj} of output r. A *referent production set* (or production possibility set) contains all input-output feasible combinations. Formally this can be stated as follows:

$$\Phi = \{ (X, Y) \mid \text{Input vector X can produce output vector Y} \} \quad (1)$$

The definition of the production possibility set is strengthened further using the following postulates:

Postulate 1. (*Non-Stochastic*) All observed operating DMU sare included in the referent set,

Postulate 2. (*Inefficiency or Free disposal*)

(a) If $(X, Y) \in \Phi$ and $X' \geq X$, then $(X', Y) \in \Phi$

(b) If $(X, Y) \in \Phi$ and $Y' \leq Y$, then $(X, Y') \in \Phi$

Postulate 3. (*Ray Unboundedness*)

If $(X, Y) \in \Phi$ then $(kX, kY) \in \Phi \; \forall \; k > 0$

Postulate 4. (*Convexity*)

If $(X^j, Y^j) \in \Phi$, $j = 1, ..., n$ and μ_j are non-negative indices

such that $\sum_{j=1}^{n} \mu_j = 1$, then $\left(\sum_{j=1}^{n} \mu_j X^j, \sum_{j=1}^{n} \mu_j Y^j\right) \in \Phi$.

Postulate 5. (*Minimality assumption*)

Φ is the intersection of all $\hat{\Phi}$ satisfying Postulates 1, 2, 3, 4 and subject to the condition that each of the observed vectors $(X^j, Y^j) \in \hat{\Phi}$, $j = 1, ..., n$.

Postulates 1, 2, 4 and 5 can be used to define a constant returns to scale (CRS) production possibility set shown below in (2).

$$\Phi_{CRS} \equiv \{ (X, Y) \in R_{+}^{m+s} | \; X \geq \sum_{j=1}^{n} \lambda_j X^j, Y \leq \sum_{j=1}^{n} \lambda_j Y^j, \lambda_j \geq 0 \} \qquad (2)$$

Exclusion of postulate 3 will lead to the definition of a variable returns production possibility set (VRS) shown below in (3).

$$\Phi_{VRS} \equiv \left\{ \begin{array}{l} (X, Y) \in R_{+}^{m+s} | \; X \geq \sum_{j=1}^{n} \mu_j X^j, Y \leq \sum_{j=1}^{n} \mu_j Y^j \\ \qquad \mu_j \geq 0 \text{ are scalars with } \sum_{j=1}^{n} \mu_j = 1 \end{array} \right\} \qquad (3)$$

The CRS and VRS production possibility sets in (2) and (3) correspondingly have in common a fundamental feature that include as members of the production possibility set *linear combinations* of inputs and outputs of observed DMUs. The convexity property, however, characterises the frontier of the production possibility set. In the CRS case the frontier is defined as a *conicall hull* whilst in the VRS case the frontier is defined as a *convex hull* of the production possibility set.

To each production possibility set there corresponds an efficient frontier which consists of a subset of its DMUs that satisfy the property of efficiency. Notice here that the concept of an efficient frontier is linked with the production possibility set. Technical efficiency can be defined as input-saving, output-augmenting or a combination of the two. In an input-saving sense the efficiency E_j^F of unit j under constant returns to scale can be defined as follows:

$$E_j^F = \min \{ \theta \mid X \leq \theta X^j, Y \geq Y^j \text{ and } (X, Y) \in \Phi_{CRS} \} \quad (4)$$

A production possibility (X^j, Y^j) with a $\theta = 1$ is called Farrell-efficient DMU and constitutes the Farrell efficient frontier. This definition, however, is not sufficient for defining "truly" efficient frontiersin the *Pareto* sense. Koopmans (1951) defined technical efficiency as follows:

> A producer is technically efficient if an increase in any output requires a reduction in at least one other output or an increase in at least one input, and if a reduction in any input requires an increase in at least one other input or a decrease in at least one output.
>
> (Koopmans, 1951, pp. 60]

The mathematical expression of this definition is as follows:

$$E_j^{P-K} = \max \left\{ s + d \;\middle|\; \begin{array}{l} X \leq X_j - s, \; Y \geq Y_j + d \text{ and } (X, Y) \in \Phi_{CRS} \\ s \in R_+^m \;, d \in R_+^s \end{array} \right\} \quad (5)$$

An optimal solution of s*+ d* = 0 in model (5) indicates that the corresponding assessed unit j is *Pareto-Koopmans* efficient unit. DMUs satisfying this criterion constitute an efficient subset of the frontier of a production possibility set. (In the remainder of the paper the term efficient frontier will always correspond to DMUs that satisfy the *Pareto-Koompans* criterion).

Farrell's efficiency in (4) is based on the radial contraction factor θ which does imply that at the boundary for some individual inputs (outputs) there is no scope for further reduction (expansion). Koopmans efficiency in (5) investigates the performance of each input and output of assessed DMUs beyond the radial contraction factor Θ. Let us illustrate the distinction between Farrell and Koopmans frontiers graphically using a two input production possibility set standardised per unit of output as exhibited in Figure 3.

In Figure 3 is assumed that the theoretical production function is an unknown and therefore efficiency is estimated on the basis of an *empirical* *production function* based on the performance of DMUs A, B and C. The technical efficiency of DMU P is defined as the ratio OP'/OP and it estimates the proportionate excess use of input 1 and input 2 in producing one unit of output. In the case of DMU D (and each unit on its horizontal expansion) the Farrell test will give an efficiency equal to one as the OD ray from the origin meets DMU D without any interference from the efficient frontier. Is DMU D, then an efficient one? Clearly not as DMU C uses that same amount of input 2 and less amount of input 1 to produce one unit of output. Farrell (1957) in his work appreciated the problem caused by this type of DMUs which he called "DMUs at unfinity" without, however, providing any methods for identifying the true efficiency of these DMUs. Using the Koopmans definition of efficiency it is clear that DMU D is inefficient.

3.3. Linear Programming Models for Assessing Efficiency

The frontier analysis discussion has succeeded so far in providing a systematic definition of production possibility sets and their efficient frontier. The next step will be to define some type of "metric" that would enable us to projet inefficient DMUs on the efficient frontier of their production possibility set. This can be done using the linear programming models developed by Charnes, Cooper and Rhodes (1978) which operationalised and extended the earlier work by Farrell (1957).

The technical efficiency of a DMU j_0 can be obtained using the two-stage linear programming model in (6) which is listed in Table 1. The assessment of efficiency can be done using an output expansion or input contraction orientation. In the remainder of the paper, for convenience, only the first stage of this process will be stated assuming, however, that any numerical calculations for assessing efficiency require this two stage process.

Where x_{ij} is the level of i^{th} input and y_{rj} is the level of r^{th} output of the j^{th} DMU; m and s are the dimensions of the input and output space respectively and η is the number of DMUs. The solution process in (6) yields input contraction or output expansion efficiencies obtained from a two-stage process. Stage 1 seeks to identify the maximim *pro rata* imput decrease or output increase. The optimal solutions obtained correspond to the *Farrell type,* of efficiency discussed earlier. Stage 2 investigates the potential extra input reduction or output expansion beyond what is already achieved at the first stage[5]. The combined solutions from stage 1 and stage 2 can be used for identifying *Pareto-Koopmans* efficient units.

A DMU j_o will be efficient if and only if the solution of the input contraction LP model is $\theta^* = 1$ and $s_i^{-*} = s_r^{-*} = 0 \ \forall \ i, r$. Similarly DMU j_o will be efficient if and only if the solution of the output expansion LP is $z^* = 1$ and $s_i^{-*} = s_r^{-*} = 0 \ \forall \ i, r$.

Charnes et al. (1978) baptised the method used for assessing efficiency "data envelopment analysis" (DEA) in an attempt to describe the rationele of the method: use the relatively efficient DMUs of a production possibility set in order to create an *efficient envelope* for inefficient DMUs. From the solution of model (6) emanate a number of observations summarised next.

- Model (6) assesses the efficiency of DMUs under constant returns to scale by solving an LP problem for each observed DMU in the production possibility set Φ_{CRS}.
- The model adopts an *Offebsive* mechanism for assessing performance. That is for each assessed DMU j_o the solution process seeks to identify a comparator (or combination of) efficient DMU $\left(\sum_{j=1}^{n} \lambda_j^* x_{ij}, \sum_{j=1}^{n} \lambda_j^* y_{rj}\right)$ that dominate j_o in all input/output dimensions.
- There is an inverse relationship between the input contraction θ^* and output expansion z^* efficiencies under constant returns to scale. It can be proven (see Seiford and Thrall (1990)) that in the optimal solution of (6) the following relation holds: $\theta^* = {}^1/_{z^*}$. This, however, should be seen as a special case (constant returns to scale) and not as a general rule.

3.4. An Alternative Formulation of DEA: The Defensive LP Model

The mathematical programming models employed in (6) were interpreted as "offensive" DEA models due to the use of composite DMUs $\left(\sum_j \lambda_j^* X_j, \sum_j \lambda_j^* Y_j, \right)$ as comparators to inefficient DMUs. An alternative (value based) formulation can be given, however, for assessing efficiency based on the dual form of the models in (6). Value based models will present DEA in the light of a generalised total factor productivity index often met in accounting and economic literature.

A total factor productivity index TFP of a DMU or firm can be defined as the weighted sum of its outputs divided by the weighted sum of its inputs. Using the notation used earlier in DEA TFP_j of unit j is given in (7).

$$TFP_j = \frac{\sum_{r=1}^{s} u_r' \, y_{rj}}{\sum_{i=1}^{m} v_i' \, x_{ij}} \quad ; u_r', v_i' \geq 0 \tag{7}$$

The selection of the weights for inputs v_i' and outputs u_r' respectively in (7) leads to the value of TFP for individual firms. In the absence of market prices so that the TFP can be converted into monetary terms one needs to assign arbitrary weights reflecting the relative importance of individual inputs/outputs in assessing productivity. Very often an assignment of equal weights among inputs and outputs is used to resolve the problem of input/output aggregation. The presence of non-commensurate inputs/outputs causes extra difficulties in the assessment of weights of importance.

In DEA, the TFP formulation is enhanced by selecting weights based on a comparative basis. In other words, the weights are treated as variables of an optimisation problem that seeks to maximise the TFP of indicidual DMUs subject to the constraint that no other unit can achieve a TFP value higher than unity (or some other upper limit). This is similar to the engineering definition of efficiency where the energy produced by a process cannot exceed the energy consumed for its generation, Charnes et al. (1985).

The mathematical formulation of this model is given in (8) as was developed by Charnes et al. (1978).

$$\begin{aligned} \max_{v_i', u_r'} \quad & \frac{\sum_{r=1}^{s} u_r' \, y_{rjo}}{\sum_{i=1}^{m} v_i' \, x_{ijo}} \\ \text{s.t.} \quad & \frac{\sum_{r=1}^{s} u_r' \, y_{rj}}{\sum_{i=1}^{m} v_i' \, x_{ij}} \leq 1 \quad \forall j = 1, \ldots, n \\ & \frac{u_r'}{\sum_{i=1}^{m} v_i' \, x_{ij}} > 0 \quad \forall r \\ & \frac{v_i'}{\sum_{i=1}^{m} v_i' \, x_{ij}} > 0 \quad \forall i \end{aligned} \tag{8}$$

The model in (8) is a linear fractional programming problem which can be converted to an ordinary linear programme using the Charnes and Cooper (1961) transformation. However, the important feature of this model lies more on the interpretation of its mechanism rather than on its mathematical transformation.

> An assessed DMU j_c "chooses" the set of weights ($vi^{'JO}$, $u/^{JO}$) that maximise its efficiency TFPj. The same weights are then attached to all other DMUs which try to "defend" their efficiency. If no other DMU reaches a higher efficiency score using the weights of the assessed DMU jo the DMU is efficient; otherwise inefficient.

Based on this rationale the model (8) will be called a *defensive* DEA model. The linear programming equivalent (for the output expansion case[6]) of model (8) is provided next in (9) [see Table 2] which is the dual mathematical model for the output expansion in (6).

The solution obtained via the offensive and defensive DEA models are linked via the duality theorem in mathematical programming. Therefore, they yield the same objective function value whilst their variables are linked via the Strong Complementary Slackness Condition.

3.5. Decomposing Technical Efficiency

The technical efficiency obtained by (6) is under constant returns to scale. Banker et al. (1984) relaxed this assumption and developed ways for disentangling technical efficiency into *scale* and *pure technical components.* The idea is illustrated using the small numerical example in Figure 4.

Figure 4 represents a single input-output production technology made of DMUs U1-U6. Under the assumption of constant returns to scale unit U2 is the only efficient DMU as it has the highest output per unit of input (5/3). The efficient frontier in this case is made of the conical hull OU2E which is an envelopment surface that can be stated as $\{\lambda\ (X2,\ Y2)\ |\ \lambda > 0\ \}$. DMU U5, therefore should expand its output by an amount of U5E or contract its input by an amount of U5G in order to be technically efficient.

The frontier is developed under the assumption that DMU U2 can be extrapolated to points, say, E and G without altering its output to input ratio. Relaxing this assumption one may redefine the efficient frontier without allowing scale extrapolations. The best observed practices, therefore, will be selected on the basis of performance given their scale of operation. The frontier in Figure 4 will be redefined, therefore, to be the piece-wise line U1U2U3U4. This frontier will be called a variable returns to scale (VRS) frontier and is made of convex combinations of the extreme points lying on its surface.

DMU U5 is an output-inefficient unit projected on the envelopment surface U2U3 defined as $\{\mu_2 (X_2, Y_2) + \mu_3 (X_3, Ys) I \mu_2 + \mu_3 = 1 \}$. DMU U5, in this case, should expand its output by a factor of 2.10 which is equivalent of the segment U5F (47.5% efficiency). In the input side DMU U5 should contract its input by a factor of 1/2.155 which is equivalent of the segment U5K (46.4% efficiency).

There is a number of important observations emanating from the VRS frontier.

- The orientation of the efficiency assessment (input or output) affects the facet of the projection when the VRS assumption is made and therefore input and output efficiency of a DMU will not be the same
- Combining the constant and variable return to scale frontier we can define a new efficiency component, namely the scale efficiency of a DMU. For example the output scale efficiency of DMU U5 is LF/LE.

Finally a frontier of mixed character can be developed where extrapolations are permitted for only a subset of efficient DMUs. Let us consider the piece wise segmant OU2U3U4. This will be defined as a non-increasing returns to scale (NIRS) frontier. Under this assumption, the scale size of technical efficient units, e.g. unit U2, can be extrapolated for comparisons with smaller, e.g. unit U1, but not larger units, e.g. unit U3. This type of frontier is used very rarely in the DEA literature, Tulkens et al. (1993), Fare et al. (1985).

Banker et al. (1984) and Fare et al. (1985) extended the original DEA models in order to estimate efficiency under the new set of assumptions. The offensive and defensive (dual) version of these models for an output expansion case are provided in (10) [see Table 3].

Model (10) differs from the original DEA model in (6) in that it has an extra (convexity) constraint in the offensive model and an extra free variable (ω) in the defensive model. The changes for the non-increasing returns to scale are also provided in the last row of the formulation.

3.6. Economies of Scale

The definition of DEA efficient frontiers has been associated with scale related issues. As a result efficient frontiers that satisfy three different assumptions of returns to scale were developed. A constant returns to scale frontier assumes that proportionate input reductions (increases) would be followed by equiproportionate output reductions (increases). A variable returns to scale assumption allows deviations in both directions. These directions constitute the nature of scale inefficiency and are listed below.

- A DMU operates under local increasing returns to scale if a proportionate increase (decrease) to its inputs will result in a higher than proportionate increase (decrease) to its outputs.
- A DMU operates under local decreasing returns to scale if a proportionate increase (decrease) to its inputs will result in a lower than proportionate increase (decrease) to its outputs.

The discussion will be facilitated using the geometric illustration of Figure 5.

As discussed earlier DMU U2 has the highest ratio of output per unit of input and therefore is the only efficient DMU under an assumption of constant returns to scale. A different efficient frontier is obtained, however, under a variable returns to scale assumption. DMUs Ul, U3 and U4, therefore, are pure managerially efficient but scale inefficient DMUs.

Banker et al. (1984) observed that the point of intersection between the constant and variable returns to scale frontiers can be used for characterising the nature of scale inefficiencies for individual DMUs. The segments below DMU U2 characterise local increasing returns to scale whilst the segments above DMU U2 characterise local decreasing returns to scale.

A numerical criterion for characterising increasing or decreasing returns to scale is the scale indicator A which gives an estimate of the extent to which CRS efficient DMUs adjust their scale size to be compared with inefficient units. In the case of DMU U5 the scale indicator A (output expansion) can be defined as $A^{\circ} = OE/OU2$. As $A^{\circ} > 1$ this implies that the nonoptimal scale DMU U5 is larger than the scale of DMU U2 which operates under constant returns to scale. For the input contraction orientation of DMU U5 the ratio $A^{1} = OG/OU2$ is less than unity which implies a non-optimal scale lower than DMUs' U5 unit scale.

To characterise economies of scale in a multi-input multi-output case Banker et al. (1984) gave a set of criteria which were generalised latter by Banker and Thrall (1992). A different set of criteria have also been suggested by Fare et al. (1985). It is argued that the question of identifying economies of scale using DEA needs further elaboration for comparing and integrating the alternative tests suggested in the literature.

The criteria for characterising returns to scale as developed by Banker and Thrall (1992) are listed in Table 4.

The criteria for characterising economies of scale in DEA are based on the solutions of the offensive DEA model under constant returns to scale and/or the defensive DEA model under variable returns to scale (output expansion). The scale factor A has already been discussed after Figure 5. Economies of scale can also be characterised using the sign of the ω variable estimated by the solution of the defensive VRS model. For example, a DMU with $\omega > 0$ in output expansion efficiency operates under local decreasing returns to scale whilst in an input contraction efficiency the criterion operates with reverse signs and therefore $\omega > 0$ characterises local increasing returns to scale. In Figure 5 the ω variables denoted as ω^{1} and ω° respectively correspond to input contraction and output expansion cases.

We shall illustrate now the returns to scale investigation for the six DMUs used in our example. The relevant information is provided in Table 5.

Unit U2 is the only technical efficient DMU and therefore operates a most productive scale size. As technical efficient (CRS) DMUs operate under constant returns to scale possible estimation of ω values has not any scale efficiency relevance. For technically inefficient DMUs, however, the estimation of the range of ω values is essential for characterising the presence of economies of scale. It is interesting to observe from Table 5 that for *managenally efficient but* scale inefficient DMUs the ω variable takes multiple optimal values.

4. Recent Developments of Frontier Analysis

Frontier analysis has witnessed considerable expansion during the last twenty years as an econometric and operational research method. Seiford (1993), provides a comprehensive listing of most of the published and unpublished frontier analysis literature. The number of 500 papers listed in Seiford's literature review indicates the widespread theoretical and applied expansion of the field. The latter, however, makes it difficult to keep up with this expansion in a review paper. The review is organised by thematic area of development without following a chronological time progress.

4.1. The Structure of the Production Possibility Set

This is perhaps the area where most research effort has been concentrated since the original development of DEA by Charmes et al. (1978). The first extension by Banker et al. (1984) relaxed the constant returns to scale production possibility set to one of variable returns to scale. Banker and Morey (1986a) moved further by making distinctions between controllable and uncontrollable inputs and outputs in assessing technical efficiency. Banker and Morey (1986b) also introduced categorical variables in assigning priorities in the comparisons between DMUs that satisfy given properties. For example, in assessing the performance of restaurants one may restrict efficient restaurants that have drive-in facilities to be compared with restaurants without these facilities but not the other way round. This idea, was generalised later by Dyson et al. (1993) introducing the notion of *multiple production functions* within a given production possibility set.

The convexity of a production possibility set was relaxed as early as (1983) by Deprins et al. in assessing the performance of post offices in Belgium. This extension received wide publicity and was named Free Disposal Hull (FDH) about seven years later with Tulkens and his associates at the Centre of Operations Research and Econometrics leading towards that research direction. The FDH idea is based on the observation that the production possibility set should be made by firms using inputs and outputs without however recognising linear combinations of the observed firms as members of the production possibility set. The FDH efficient frontier is illustrated graphically in Figure 6 using our earlier numerical example.

Under the FDH assumption the efficient frontier is determined by Ul, U2, U3, U4 and has a step-wise shape. The output expansion required for DMU U5 is now U5C as compared to U5B under VRS and U5A under CRS assumptions respectively. DMU U5 is inefficient as compared with only DMU U2. This is because U5 is located within the area of "dominance" of DMU U2 (shaded area) which assumes that the efficient DMU U2 will always deliver its current outputs if it is provided with more input (disposability). The FDH idea was extended further by Athanassopoulos and Storbeck (1995) in assessing spatial efficiency.

4.2. Defining Efficiency Metrics

Earlier developments of frontier analysis made clear that the definition of production possibility sets and their efficient frontiers should be kept separate from the assessment of the efficiency of DMUs. To do this one needs to employ "metric functions" that measure the distance between inefficient DMUs and the efficient frontier. A variety of metric measures followed the initial developments by Debreu (1951) and Farrell (1957). One needs to mention the work by Fare and Lovell (1978) which developed the so-called *Russell* efficiency index; the additive DEA model developed by Charnes et al (1985) and the Fare et al. (1985) and (1993) attempt to define hyperbolic efficiency metrics.

In summary, the main debate on efficiency measures is mainly focused on whether they have *radial* on *non-radial* nature. Economists like Russell (1985) express the view that efficiency indices should be homogenous of degree-1 and as non-radial measures fail to satisfy this test they are "undesirable". In the management science field Thanassoulis and Dyson (1992) have found advantages in using non-radial efficiency measures for target setting.

4.3. Weight Restrictions and Value Frontiers

The original development of DEA by Charmes et al (1978) was based on the assumption that each assessed unit should have free choice in selecting weights for inputs and outputs without any preliminary restrictions (see defensive DEA model in (9)). However, this era lasted until 1986 when Thomson et al. (1986) argued that in selecting potential sites for locating a nuclear research laboratory they had to restrict the flexibility of weights in order to reduce the number of DMUs assessed as efficient by the standard DEA. This attempt was followed by a very rapid expansion of ideas on how weight constraints should be imposed in assessing efficiency, Dyson and Thanassoulis (1988), Charnes et al (1990) and Beasley (1988), Thanassoulis et al. (1994), Dyson et al (1993) and Cook et al (1990).

Dyson et al (1993) in a later attempt sought to investigate the consequences of the use of weight restrictions on the production possibility set of units, the efficient frontier and finally the efficiency metrics. Weight restrictions have so far been used either for reducing the number of efficient DMUs in ordinary DEA or for incorporating experts' opinion of the importance of some of the inputs/outputs on the assessed efficiency. These issues are, however, subject to intense research focus seeking to develop systematic methods for setting weight restrictions and understanding the full impact of weight restrictions on the efficiency process.

4.4. Integrating the Time Element Into Efficiency Assessments

The original use of frontier analysis was based on cross section observations and therefore the efficiency of DMUs was assessed for a particular time period. Charnes et al. (1985) introduced the notion of *window-analysis* for assessing performance over time. In window analysis, DMUs from adjacent[8] time periods were combined into clusters and their efficiency was assessed. Then a sequence of other clusters were created and assessed by removing the earlier and introducing some later time periods into the assessed clusters.

The major integration of time into the assessment of efficiency, however, came later by Fare et al (1992) which introduced the so-called *malmquist* indices for assessing efficiency variations over time. In the Malmquist analysis a combination of four performance indices are estimated for each DMU. For example if the unit is observed in time period t and t+1 then we estimate the efficiency of the observed unit at t against the frontiers at t and t+1; we also estimate the efficiency of the observed unit at t+1 against the frontiers at t and t+1. Fare et al. (1992) use the geometrical mean of these indices to define the technological progress/regress of DMUs. *Malmquist* analysis is an area of rapid research expansion during the last four years. Attempts to use Malmquist types of analysis for decision support can be found in Athanassopoulos and Thanassoulis (1994).

4.5. Organisational Science, Frontier Analysis and Decision Support

Lewin and Minton (1986) launched an attempt to open a research debate for opening a communication network between the general managerial concepts of performance and organisational effectiveness and frontier analysis. Epstein and Henderson (1988) moved further and examined the appropriateness of frontier analysis as tools for *control* and diagnosis. At the operational level, Lewin et al. (1993) used frontier analysis to support the identification of strategic groups in the brewing industry in the USA. Athanassopoulos and Ballantine (1995) on the other hand worked at the corporate firm level to examine the performance differences between strategic groups in the UK grocery industry using frontier analysis as the perfosmance yardstick.

4.6. The Econometric School of Frontier Analysis

The early development of non-parametric frontier analysis in (1978) was followed by a series of studies comparing DEA with traditional econometric techniques. Banker et al. (1988) used a known production function with simulated observations to compare econometric with envelopment frontier analysis. More recently Thanassoulis (1993) compares ordinary least squares regression with DEA as tools for performance measurement and target setting. These studies sought to emphasise the differences between econometrics and envelopment frontier analysis and resulted some times in considerable methodological debates, (see the debate between Charnes, Cooper and Sueyoshi (1988), and Evans and Heckman (1988) about the break-up of Bell telecommunications in the USA).

The second school seeks to relax the deterministic nature of DEA in at least two dimensions. Land et al (1991) and (1992) sought to estimate efficiency by introducing uncertainty into the coefficients of the DEA assessment. This study was based on the chance-constraint approaches developed long ago by Charnes and Cooper (1959).

4.7. Computational Aspects of Frontier Analysis

The non parametric frontier analysis methods that were discussed in this chapter have a linear programming nature. Thus the computational problems that emerge from frontier analysis can be addressed using the powerful linear programming codes that are commercially available (e.g. GAMS, XPRESS-MP, SAS/OR, LP88, AIMMS). At the time this paper is written there are two commercial software specifically designed for non-parametric frontier analysis. The first, is called IDEAS and developed by Ali (1989) whilst the second in called the WarwiclcDEA (1987) and has been developed by the DEA research team at the University of Warwick. A numerical illustration of the computational efficiency of the Warwick-DEA is given in Table 6.

5. Conclusion

The assessment of performance of economic systems has undoubtelly gained substantial publicity over the last two decades. Traditional ways of management and decision making are constantly revised and ineviatably new decision support tools are necessary to facilitate the new managerial approaches. Traditional mechanisms for control and performance diagnosis have an accounting bias which do not provide all necessary information concerning the assessment of performance. Frontier analysis seeks to co-ordinate and integrate research efforts from different disciplines in assessing the productive efficiency of DMUs. The rapid expansion of the field during the last twenty years indicates healthy prospects of the method in assessing performance. The frontiers of frontier analysis are constantly expanded with applications in previously "virgin" research areas. These areas include *inter alia* the general problem of provision of decision support, the linkage between resource allocation and performance measurement, the methodological progress in areas such as weights restrictions and target setting and finally the technical progress towards developing graphical interfaces for presenting the performance assessment results to the non-technical audience.

Footnotes

1. In the 1988-1993 recession in the UK most organisations sought to increase their performance in order to survive under the adverse market conditions.

2. Hoagland (1964), however, argues that many of Taylor's theories can be found published in previous research work.

3. Shareholders in the public sector are assumed to be the taxpayers.

4. Increasing (decreasing) the inputs by a constant factor would increase (decrease) the outputs with same factor.

5. The original formulation by Charnes et al. (1978), compounds the two stages in one stage by including the slack variables in the objective function of the first stage multiplied by very small coefficients. This method despite its appeal in the literature creates computational difficulties described by Ali and Seiford (1989).

6. An equivelent formulation holds for the input contraction case which was omitted to avoid repetitions.

7. Units operating under constant returns to scale are scale efficient, Banker (1984), and therefore the criteria do not have any relevance.

8. The number of time periods combined was decided arbitrarily.

Appendix

TABLE 1

Linear Programming for Assessing Technical Efficiency (6)

CRS^I - Input contraction	
Stage 1 - Contraction factor (θ)	Stage 2 - Pareto test (s_j, s_r)
$\min_{\theta, \lambda_j'} \ \theta^* = \theta$	$\min_{s_i, s_r, \lambda_j'} \ \sum_{i=1}^{m} s_i^- - \sum_{r=1}^{s} s_r^+$
$\sum_{j=1}^{n} \lambda_j' x_{ij} \leq \theta x_{ij_o} \quad \forall i$	$\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = \theta^* x_{ij_o} \quad \forall i$
$\sum_{j=1}^{n} \lambda_j' y_{rj} \geq y_{rj_o} \quad \forall r$	$\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{rj_o} \quad \forall r$
$\lambda_j' \geq 0$, θ free	$s_i^-, s_r^+, \lambda_j \geq 0$
CRS^0 - Output expansion	
Stage 1 - Expansion factor (z)	Stage 2 - Pareto test (d_j, d_r)
$\max_{z, \lambda_j'} \ z^* = z$	$\min_{d_i^-, d_r^+, \lambda_j} \ -\sum_{i=1}^{m} d_i^- - \sum_{r=1}^{s} d_r^+$
$\sum_{j=1}^{n} \lambda_j' x_{ij} \leq x_{ij_o} \quad \forall i$	$\sum_{j=1}^{n} \lambda_j x_{ij} + d_i^- = x_{ijO} \quad \forall i$
$\sum_{j=1}^{n} \lambda_j' y_{rj} \geq z y_{rj_o} \quad \forall r$	$\sum_{j=1}^{n} \lambda_j y_{rj} - d_r^+ = z^* y_{rjO} \quad \forall r$
$\lambda_j' \geq 0$, z free	$d_i^-, d_r^+, \lambda_j \geq 0$

TABLE 2

Offencive - Defensive Output Expansion DEA Models (9)

CRS^O - Offensive	CRS^O - Defensive
$\max_{\lambda_j, z} \ z$	$\min_{v_i, u_r} \ E^O_{CRS} = \sum_{i=1}^{m} v_i x_{ij_o}$
$-\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = -x_{ij_o} \ i = 1, \ldots, m$	s. t $\quad \sum_{r=1}^{s} u_r y_{ij_o} = 1$
$\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = z y_{rj_o} \quad r = 1, \ldots, s$	$\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} - t_j = 0$
z free and $\lambda_j \geq 0$, $\forall j$ $s_i^-, s_r^+ \geq 0$	$v_i, u_r, t_j \geq 0$

TABLE 3
DEA Models for Pure Technical Efficiency (10)

Offensive output expansion **variable returns to scale (VRS)**	**Defensive output expansion** **variable returns to scale (VRS)**
$\max_{\mu_j, \rho} \ \rho$ $\sum_{j=1}^{n} \mu_j x_{ij} \leq x_{ij_o} \quad i = 1, \ldots, m$ $-\sum_{j=1}^{n} \mu_j y_{rj} \leq -\rho y_{rj_o} \quad r = 1, \ldots, s$ $\sum_{j=1}^{n} \mu_j = 1$ ρ free and $\mu_j \geq 0, \forall j$	$\min_{v_i, u_r} \ E^{O}_{VRS} = \sum_{i=1}^{m} v_i x_{ij_o} + \omega$ s. t. $\sum_{r=1}^{s} u_r y_{rj_o} = 1$ $\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} + \omega \geq 0$ $v_i, u_r \geq 0$; ω free
Offensive output expansion **non-increasing returns to scale NIRS**	**Defensive output expansion** **non-increasing returns to scale NIRS**
In above formulation change $\sum_{j=1}^{n} \mu_j \leq 1$	In above formulation change $\omega \geq 0$

TABLE 4
Criteria for Identifying Returns to Scale Output Expansion Case

Characterisation	**Offensive model (6)**	**Defensive model (10)**
Local increasing return to scale	$\Lambda = \sum_{j=1}^{n} \lambda_j^* < 1$ for all optimal solutions	$\omega^{min} < \omega^{max} < 0$
Local decreasing returns to scale	$\Lambda = \sum_{j=1}^{n} \lambda_j^* > 1$ for all optimal solutions	$0 < \omega^{min} < \omega^{max}$
Constant returns to scale	$\Lambda = \sum_{j=1}^{n} \lambda_j^* = 1$ for all optimal solutions	—[7]

TABLE 5
Returns to Scale for the Six DMU
(Output expansion - Input contraction)

UNIT	Model (6) $\sum_{j=1}^{n} \lambda_j^*$ Output	Input	Model (10) $\omega^{max} \leq \omega \leq \omega^{min}$ Output	Input	Returns to scale Output	Input
U1	0.6	**0.4**	(-∝, -2]	[0.66, 1]	Incr	Incr
U2	1	**1.0**	-	-	Con	Con
U3	2	**1.4**	[0.42, 0.71]	[-2.5, 0.75]	Decr	Decr
U4	3	**1.6**	[0.62, 1]	(-∝, -1.6]	Decr	Decr
U5	1.6	**0.6**	1	0.26	Decr	Incr
U6	1.3	**0.2**	3	0.50	Decr	Incr

TABLE 6
Efficient Algorithmic Procedures Using the Warwick-DEA
(3 Inputs - 3 Outputs in 486DX processor)

No. of DMUs	Efficient DMUs 10%	Efficient DMUs 20%
50	4 sec	9 sec
200	10 sec	24 sec
600	125 sec	220 sec
2000	600 sec	1020 sec

Figure 1
Farrell Decomposition of Efficiency

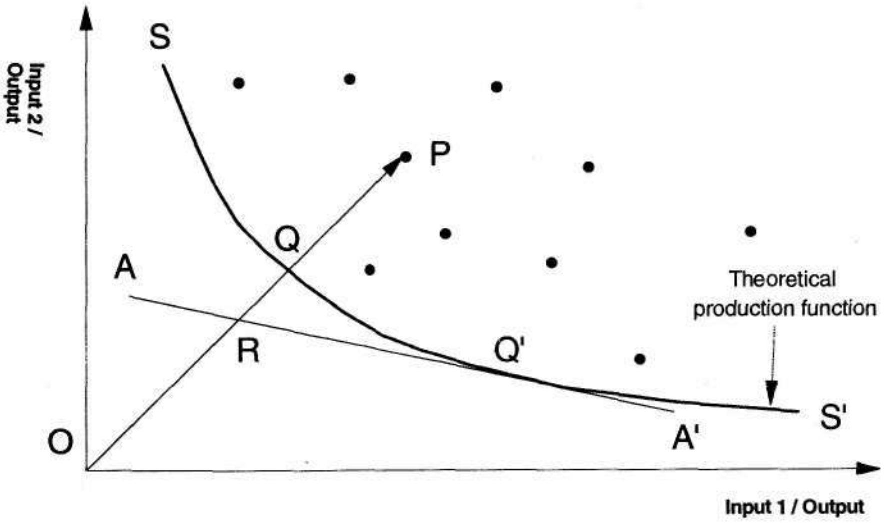


Figure 2
Input - Output Systems for Frontier Analysis

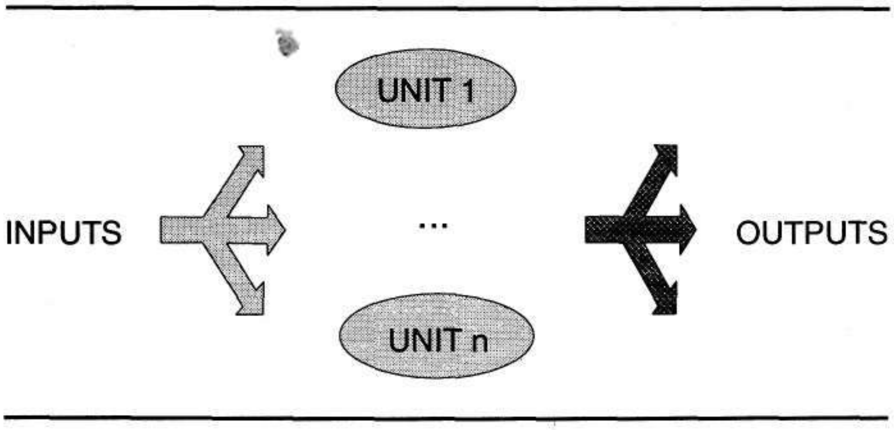


Figure 3
Farrell Vs. Pareto Efficient Frontiers

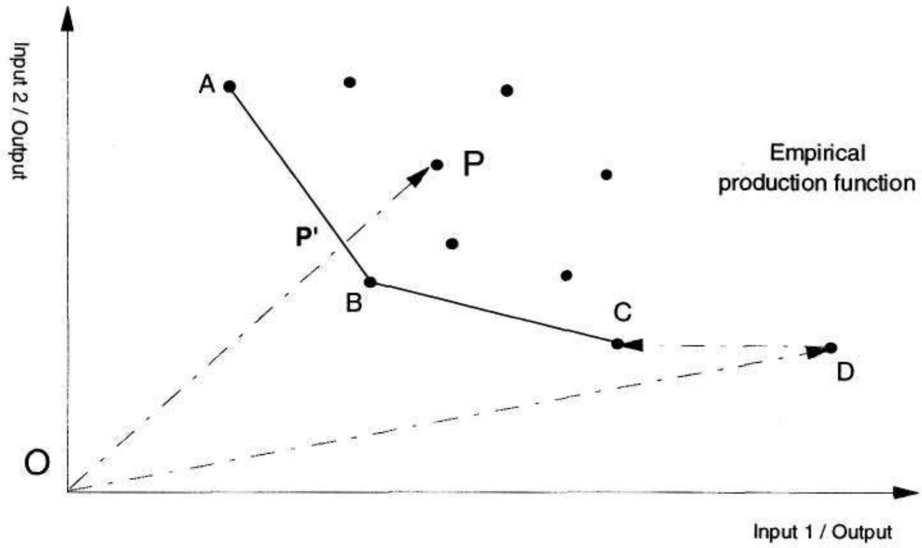


Figure 4
Efficient Frontiers and Returns to Scale Assumptions

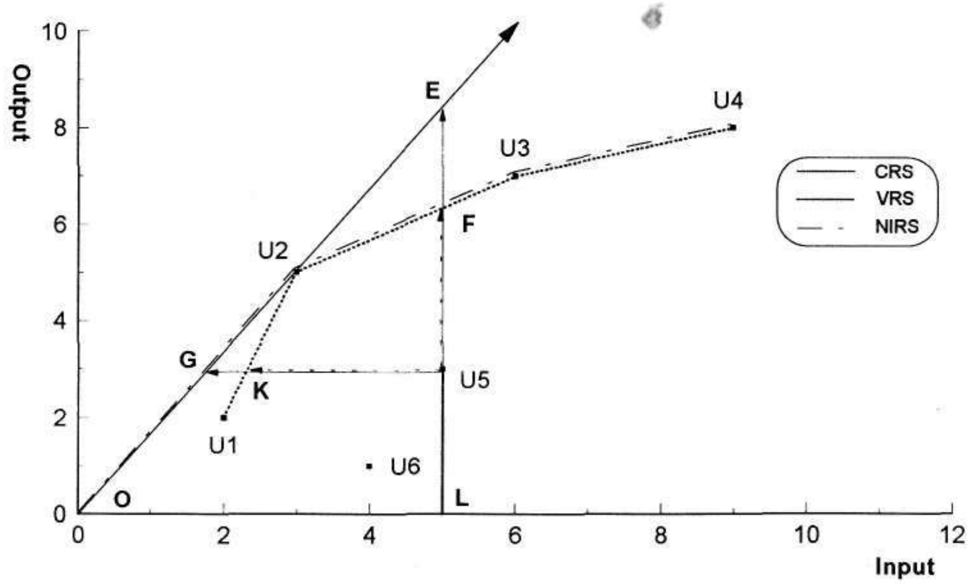


Figure 5
Classifying Local Economies of Scale

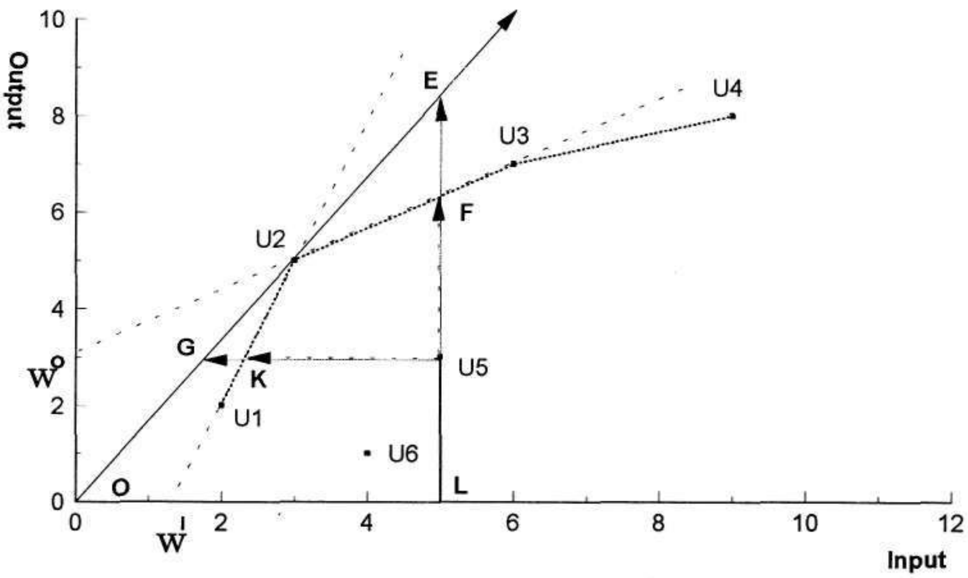
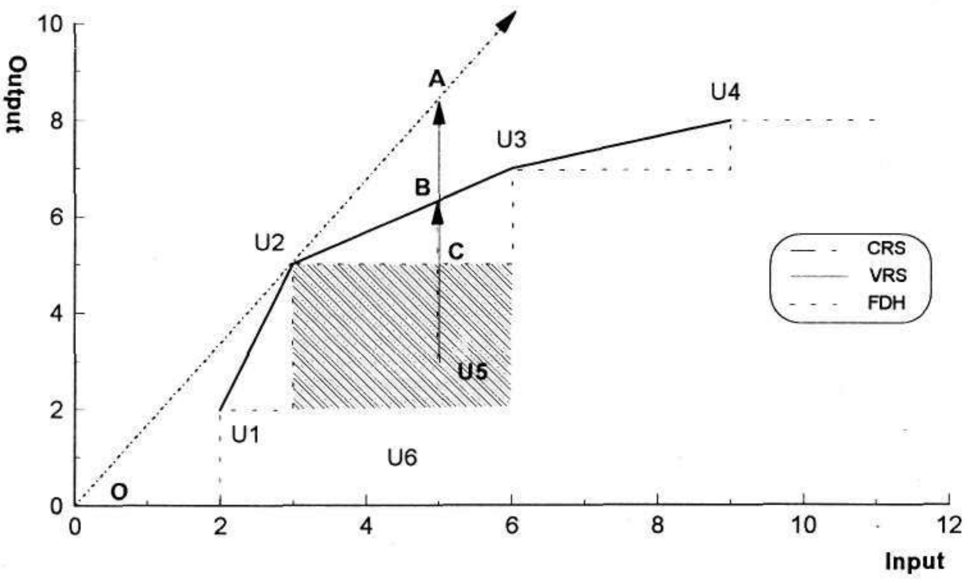


Figure 6
Free Disposal Efficient Frontiers



References

Aigner D., K. Lovell and P. Schmidt (1977), "Formulation and estimation of stochastic frontier production function models", Journal of Econometrics, 6 (1), pp. 21-37.

Aigner D. and S., Chu, (1968), "On estimating the industry production function", American Economic Review 58 (4), pp. 826-839.

AH A. and L. Seiford, (1989), "Computational accuracy and infinitesimals in data evnelopment analysis", Technical report, University of Massachusetts at Amherst, Amherst.

AHA. andL. Seiford, (1993), "The mathematical programming approach to efficiency analysis", In Fried H., K. Lovell, and S. Schmidt, The measurment of productive efficiency: Techniques and applications, Oxford University Press, New York.

Athanassopoulos, A., and E., Thanassoulis, (1991), "Separating market efficiency from profitability and its implications for planning", Journal of Operational Research Society, 46 (I), pp. 20-35.

Athanassopoulos, A. andJ., Ballantine, (1995), "Ratio and frontier analysis for assessing corporate performance: Evidence from the Grocery industry in the UK", Operational Journal of the Research Society, 46 (4), pp. 427-440.

Athanassopoulos A. and E. Thanassoulis, (1994), "Assessing the marginal impacts of investments on the performance of organisational units" Internatrional Journal of Production economics (In press).

Athanassopoulos A. and J. Storbeck, (1995), "Non-parametric models for Spatial Efficiency", Journal of Productivity Analysis, 6 (2).

Athanassopoulos A. and N. Tatsos, (1992), "Econometric modelling as an aid for reorganisation policies: The case of the Greek local authorities", The Institute of Applied Econometrics, Econometrics Europe 2000, Belgium, Brussels, Published proceedings.

Banker R., A. Charmes and W. Cooper, (1984), "Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis", Management Science, Vol. 30 (9), pp. 1078-1092.

Banker R., (1984), "Estimating most productive scale size using data envelopment analysis", European Journal of Operational Research, Vol. 17, pp. 35-44.

Banker R. and R. Morey, (1986a), 'The use of categorical variables in data envelopment analysis", Management Sciences, Vol. 32, pp. 1613-1627.

Banker R. and R. Morey, (1986b), "Efficiency analysis for exogenously fixed inputs and outputs", Operations Research, Vol. 34, pp. 513-521.

Banker R. and R. Thrall, (1992), "Estimation of returns to scale using data envelopment analysis", European Journal of Operational Research, Vol. 62, pp. 74-84.

Beasley, J., (1990), "Comparing university departments", Omega International Journal of Management Science, Vol. 18, No. 2, pp. 171-183.

Booth P., and Cocks N.,(\990), "critical research issues in accounting standard setting", Journal of Business Finance and Accounting, pp. 511-528.

Boussofiane A., R. G. Dyson and E. Thanassoulis, (1991), "Applied data envelopment analysis", European Journal of Operational Research, Vol. 52, pp. 1-15.

Chambers R., (1990), "Applied production analysis; a dual approach", Cambridge University Press.

Charmes A., C. Clark, W. Cooper and B. Golaby, (1985), "A development study of data envelopment analysis in measuring the efficiency of maintenance DMUs in the U.S. air forces", Annals of Operations Research 2, pp. 95-112.

Charmes A., W. Cooper and E. Rhodes, (1978), "Measuring the efficiency of decision making units", European Journal of Operational Research, Vol. 2 (6), pp. 429-444.

Charmes A., W. Cooper and E. Rhodes, (1981), "Evaluating program and managerial efficiency: An application of data envelopment analysis to program follow through", Management Science, Vol. 27, pp. 668-697.

Charmes A. W., Cooper, B. Golany, L. Seiford and J. Stutz, (1985), "Foundations of data envelopment analysis for Pareto-Koopmans efficient empirical production functions", Journal of econometrics, Vol. 30, pp. 91-107.

Charmes A., W. Cooper, Z. Huang and D. Sun, (1990), "Polyhedral cone-ratio DEA models with an illustrative application to large commercial banks", Journal of Econometrics, Vol. 46, pp. 73-91.

Charmes A., W. Cooper and T. Sueyoshi, (1988), "A goal programming /constrained regression review of the Bell system break-up", Management Science, 34 (1), pp. 1-26.

Charmes A. and Cooper W., (1978), "Managerial Economics Past, Present and Future", Journal of Enterprise Management, Vol. 1, pp. 5-23.

Charmes A. and W. Cooper, (1959), "Chance-Constrained Programming", Management Science, 15, pp. 73-79.

Charmes A. and W. Cooper (1985), "Preface to topics in data envelopment analysis", Annals of Operations Research, Vol. 2, pp. 59-94.

Day A. and A. Lewin, (1994), "Strategic leaders in the US Brewing industry: A longitudinal analysis of outliers, In: Charnes, Cooper, Lewin and Seiford, eds., Data Envelopment analysis: Theory, methodology and applications.

Debreu G. (1951), "The coefficient of resource utilisation", Econometrica, 19(3), pp. 273-292.

Deprins D., L. Simar and H. Tulkens, (1983), "Measuring labour efficiency of post offices., In Marchand, Petisteau and Tulkens, eds., The Performance of public enterprises: concepts and measurement. Amsterdam: North-Holland.

Dyson R., E. Thanassoulis and A. Athanassopoulos (1993), "Weights restrictions, value ratio frontiers, and production possibility frontiers in data envelopment analysis", Watwick Business School research papers No. 97.

Dyson R. G. andE. Thanassoulis, (1988), "Reducing weights flexibility in data envelopment analysis", Journal of the Operational Research Society, Vol. 39, pp. 563-576.

Epstein M. and J. Henderson, (1988), "Data envelopment analysis for managerial control and diagnosis", Decision Sciences, 20 (1), 90-119.

Evans D. andJ. Heckman, (1988), "Natural monopoly and the Bell system: Response to Charnes, Cooper and Sueyoshi", Management Science, 34 (1), pp. 27-38.

Fare, R., S. Grosskopf, B. Lindgren, P. Roos, (1992), "productivity changes in Swedish pharmacies 1980-1989: A non-parametric malmquist approach", The Journal of Productivity Analysis, Vol. 3, pp. 85-101.

FareR.,S. Grosskopf, andK. Lovell, (1985), The measurement of efficiency of production, Boston: Kluwer-Nijhoff publishing.

Fare R. and K. Lovell, (1978), "Measuring the technical efficiency of production", Journal of Economic Theory, 19 (1), pp. 150-162.

Farrell M., (1957), "The measurement of productive efficiency", Journal ofRoyal Statistical Society A, Vol. 120, pp. 253-281.

Forsund F. and L. Hjalmarsson, (1979), "Frontier production functions and technical progress: A study of general milk processing in Swedish dairy plants", Econometrica, 47 (4), pp. 883-900.

Frantz R., (1992), "X-efficiency and allocative efficiency: what have we learned?", American Economic Review, 82(2), pp. 434-438.

Fried H., K. Lovell, and S. Schmidt, (1993), The measurement of productive efficiency: Techniques and applications, Oxford University Press, New York.

Giokas D., (1990), "Bank branch operating efficiency: A comparative application of DEA and the loglinear model", Omega International Journal of Management Science 19 (6), 549-557.

Griffiths I., (1992), "Creative accounting: How to make your profits what you want them to be", Routledgw, London.

Hogler R. and H. Hunt III, (1993), "Accounding and conceptions of control in the American corporation", British Journal of Management, Vol. 4, pp. 177-190.

Knight F., (1933), The economic organisation, New York: Harper and Row.

Koopmans (1951), "Activity analysis of production and allocation", John Wiley and Sons, New York.

Koopmans (1957), "Three essays on the state of economic science", McGraw-Hill, New York.

Land K, K. Lovell and S. Thore, (1991), "Change-constraint data envelopment analysis", IC^2 Institute discussion papers, The university of Texas at Austin.

LandK., K. Lovell and S. Thore, (1993), "Efficiency under capitalism and state socialism", Public Finance - Finances Publiques, 47, pp. 109-121.

Leibenstein H. andS. Maital, (1992), "Empirical estimation and partitioning of X-efficiency: A data envelopment approach", American economic Review, Vol. 82, pp. 428-433.

Lewin A. and Minton, (1986), "Determining organisational effectiveness: Another look, and an agenda for research", Management Science, 32 (5), pp. 514-537.

Mayston D., (1993), "Principals, agents and the economics of accountability in the new public sector, Accounting Auditing and Accountability journal, 6 (3), pp. 68-96.

RayS., (1988), "Data envelopment analysis, non-discretionary inputs and efficiency: An alternative interpretation", Socio-Economic Planning Sciences, 22 (4), pp. 167-176.

J?usse//J?., (1985), "Mesures of technical efficiency", Journal of economic theory, 35(1), pp. 109-126.

Schmidt P., (1985-1986), "Frontier production functions", Econometric reviews, 4 (2), pp. 289-328.

SeifordL. andR. Thrall, (1990), "Recent developments in Data Envelopment Analysis", Journal of Econometrics, Vol. 46, pp. 7-38.

Thanassoulis E. and Dunstand, (1993), "Guiding schools to improved performamce using data encelopment analysis: An illustration with data from a Local Education Authority", Warwick Business School research papers No. 94, (Forthcoming in the Journal of the Operational Research Society).

Thanassoulis E. andR. Dyson, (1992), "Estimating preferred targets input-output levels using Data Envelopment Analysis", European Journal of Operational research, Vol. 56, pp. 80-97.

The Warwick-DEA package (user guide), Thanassoulis, ER., (1992), Warwick Business school, University of Warwick.

Thompson R., L. Langemeier, C. Lee, E. Lee and M. Thrall, (1990) "The role of multiplier bounds in efficiency analysis with application to Kansas farming", Journal ofeconometrics, Vol. 46, pp. 73-93.

Tulkens H. and Vanden Eecjaut, (1991), "Non-frontier measures of efficiency, Progress and Regress", CORE discussion paper 9155, Centre of Operational Research and Econometrics, Universite Catholique de Louvain.

Vassiloglou M. and D. Giokas, (1990), "A study of the relative efficiency of bank branches: An application of data envelopment analysis", Journal of the Operational Research Society, 41 (7), 591-597.