# A Study in Theory and Models of Data Envelopment Analysis

***Aikaterini Kokkinou***

*Department of Economics, University of Glasgow, U.K.*

**ABSTRACT**

*Although the significance of efficient use of production resources has long been recognized, neoclassical economics assume that producers always operate efficiently. In reality, however, producers are not always efficient. This difference may be explained in terms of efficiency and some unforeseen exogenous shocks. Regarding this, parametric stochastic frontier production function, involving econometric methods, and non-parametric Data Envelopment Analysis (DEA), involving mathematical programming techniques, are the two most popular methods found in the literature of efficiency estimation. This paper aims to review the main topics related to the theoretical background of Data Envelopment Analysis, reviewing both DEA model and extensions, as well as the main characteristics, advantages and disadvantages of DEA application.*

***JEL:*** *C14, R30, O40.*

***Keywords:*** *Parametric and non-parametric frontier, Stochastic frontier analysis, Data envelopment analysis, Productive efficiency.*

## 1. INTRODUCTION

Estimation of productivity and productive efficiency is a relatively complex economic issue (for a broad overview, see Korres, 2007a). The widely used frontier approach on productive efficiency estimation makes it possible to distinguish between shifts in technology from movements towards the best-practice production frontier. Central to frontier productivity analysis is the determination of the efficient production technology, as well as the identification of those efficient decision-making units (DMUs) on the frontier and of those inefficient DMUs not on the frontier – and, for the latter, determination of the degree and sources of their inefficiency. By estimating the best-practice production function these approaches calculate technical efficiency as the distance between the frontier and the observed output. Towards this direction, different techniques have also been used to measure technical efficiency under the frontier approach. One of the most commonly used techniques is the nonparametric linear programming technique or Data Envelopment Analysis (DEA). This papers reviews these terms, providing definitions, descriptions and model forms and attempts to gives a broad background for empirical implementation.

## 2. THE DATA ENVELOPMENT ANALYSIS (DEA)

In principle, DEA uses a set of non-parametric programming techniques to identify which subset of a set of enterprises may be considered as the best[1]. DEA is, in fact, a mathematical programming approach for the construction of production frontiers and the measurement of efficiency relative to the constructed frontiers[2]. The basic idea of this approach consists of enveloping the data (the observed input-output combinations) in order to obtain an approximation of the production frontier (best-practice frontier) and using this to identify the contribution of technological change, technological catch-up, and inputs accumulation to productivity growth.

## 3. THE THEORETICAL BACKGROUND

The non-parametric Data Envelopment Analysis (DEA) approach for measuring efficiency was pioneered by Charnes, Cooper and Rhodes (1978)[3] as an attempt to overcome some of the specific weaknesses of the growth accounting approach: a particular functional form for technology, particular assumptions on market structure, and the hypothesis that markets are perfect. Charnes, Cooper, and Rhodes (1978) used mathematical programming to generalize Farrell's (1957) single-output/single-input technical efficiency measure by transforming a multiple-output/multiple-input technology into one combined output and one combined input. Charnes *et al*., (1978) described DEA methodology as a mathematical programming model applied to the observed data, providing a new way of obtaining empirical estimates of external relations such as the production functions and/or production possibility functions (Korres, 2007b).

Since then, DEA has been increasingly used in efficiency analysis in many areas especially in efficiency measurement of economic entities (e.g. Desai and Schinnar, 1987 and Sengupta, 1987). Further, Coelli and Perelman (1999) compared technical efficiency scores and they found that the choice of method does not have much influence on the results. Another example is the study by Iráizoz *et al*., (2003) who compared technical efficiency results and found correlation between the parametric and nonparametric approach. Also, Arega (2005) compares the empirical performances of the parametric distance functions and data envelopment analysis. Coelli (1995) presented a review of both parametric and non-parametric techniques used in efficiency measurement, including their limitations, strengths, and applications. Although his review indicated that parametric approaches were used more frequently than DEA, neither model appears to have dominant advantages above the other. Sharma *et al*., (1999) used Stochastic Frontier Analysis (SFA) and Data Envelopment Analysis to estimate technical efficiencies. They found that, on average, the estimated technical efficiencies were significantly higher in the SFA compared to DEA under the assumption of constant returns to scale (CRS). Under the assumption of variable returns to scale (VRS) however, the measures were quite similar. Hjalmarsson *et al*. (1996) provided results obtained from the stochastic frontier model and DEA models. Similarity and dissimilarity depended upon the inclusion of the control variables in the stochastic frontier and sequential or intertemporal specification in the DEA frontier. Johansson (2005) estimated technical, allocative, and economic input efficiency scores for an unbalanced panel, using data envelopment analysis (DEA) and stochastic frontier approach (SFA). By comparing the results it was concluded that DEA measures for technical and economic efficiency were significantly higher than the corresponding SFA measures. Serrao (2003) examined differences in agricultural productivity growth among countries and regions. Findings indicate that the mean TFP scores are higher under DEA than under SFA because DEA fits a tighter (i.e., more flexible) frontier. Hence, Serrao warned against the subjective choice of a particular approach and suggested the use and comparisons of more than one approach. DEA reports all deviations from the frontier as inefficiency, and thus should report lower efficiency scores compared to SFA. However, misspecification of the functional form by the SFA method would possibly cause lower efficiency scores relative to DEA methods.

## 4. THE SPECIFICATIONS OF THE MODEL

Data envelopment (or envelope) analysis (DEA) provides a way of measuring the relative technical efficiency of different decision making units performing the same or similar tasks. Under this assumption, DEA is used to obtain efficiency measures based on the aggregated, or 'virtual', inputs and outputs. As far as the DEA model extensions are concerned, DEA can be specified as either an output-maximizing problem or an input-minimizing problem. Input models measure efficiency in terms of the potential (proportional) reduction in input use while output models measure efficiency in term of the potential (proportional) output increase. The analysis assumes that there is a frontier technology (in the same spirit as the stochastic frontier production model) that can be described by a piecewise linear hull that envelopes the observed outcomes. Some (efficient) observations will be on the frontier while other (inefficient) individuals will be inside. The technique produces a deterministic frontier that is generated by the observed data, so by construction, some individuals are efficient. Efficiency is measured relative to the actual performance of the 'best practice' firms. By normalization, the efficiency scores range from zero to one. The best practice enterprises are given a ranking of 1 and efficiency scores are assigned to other enterprises by comparing them to the best practice enterprises. That is, solutions are sought to maximize the ratio of weighted output to weighted input for each firm (the ratio of virtual output to virtual input).

As described in McMillan and Chan (2006), let there be $n$ Decision making Units (DMUs) using varying amounts of inputs to produce outputs. There are $s$ inputs $xi$ ($i$ = 1, …, $s$) and $m$ outputs $yr$ ($r$ = 1, …, $m$). For each DMU, such as $\text{DMU}_j$ ($j$ = 1, …, $k$, …, $n$), the problem is to:

$$\max_{u,v} h_j = \frac{\sum_r u_{rj}\, y_{rj}}{\sum_i v_{ij}\, x_{ij}} \tag{1}$$

$$\text{subject to : } \frac{\sum_r u_{rj}\, y_{rj}}{\sum_i v_{ij}\, x_{ij}} \le 1 \quad \text{for} \quad j = 1, ..., n \tag{2}$$

$$u_r, v_i \ge 0$$

where $u_{rj}$ is the weight assigned each unit of output $r$ from $\text{DMU}_j$ and $v_{ij}$ is the weight assigned each unit of input $i$ used by $\text{DMU}_j$. That is, solutions are sought to maximize the ratio of weighted output to weighted input for each DMU (the ratio of virtual output to virtual input). By normalization, the efficiency scores range from zero to one. The same weights (virtual multipliers) that maximize $h_j$ for $\text{DMU}_j$ are applied to the inputs and outputs of all DMUs in the solution to the problem for $\text{DMU}_j$. This solution process is repeated for each DMU. Hence, because the weights can vary for each solution, the efficiency scores determined are those most favourable to each DMU[4].

## 5. DEA: USING AN EXTENDED APPROACH

DEA extension models can be categorized following various criteria. According to the shape of the production function, they are categorized as *constant returns to scale* DEA (CRS DEA) and *variable returns to scale* DEA (VRS DEA) models. Under CRS, the form of the envelopment surface of the constructed production frontier is a conical hull, while under VRS, it is a convex hull. CRS DEA model is often reported as CCR from Charnes, Cooper and Rhodes (1978), who first introduced it, while VRS DEA model is often report as BCC model from Banker, Charnes and Cooper (1984). Finally, according to the selected orientation, DEA models are categorized as input-oriented, output-oriented or no-oriented models. Selection of a model's orientation determines the path of inefficient DMUs to the efficient frontier. In non-oriented models, output slack and input excess are considered comparable in that neither should receive greater scrutiny than the other, while in oriented models, either inputs or outputs preempt the other in that proportional movement toward the frontier is first achieved in input or output space, respectively. Thus, in an output orientation model, the objective is to produce the maximum amount of outputs with a given set of inputs, so the efficiency frontier is constructed via proportional augmentation of all outputs. In an input orientation model, the objective is to produce the desired output with minimum amount of inputs; therefore, the efficiency frontier is constructed via proportional reduction in all inputs. Under constant returns-to-scale, the measure of input efficiency will be the inverse of the measure of output efficiency (Fare and Lovell, 1978).

### 5.1. The Constant Returns to Scale (CRS) DEA Model (Input Oriented CRS DEA)

Assume there are *N* firms in an industry. For each firm, *K* inputs and *M* outputs are represented by the column vectors, $x_i$ and $y_i$, respectively. The *KxN* input matrix, *X*, and the *MxN* output matrix, *Y*, represent the data for all firms. A rational way of introducing DEA is via the ratio of all outputs over all inputs, such as:

$$\frac{u'yi}{v'xi} \tag{3}$$

where, *u* is *Mx*1 vector of output weights and *v* is a *Kx*1 vector of input weights. The optimal weights are obtained by solving the mathematical programming problem:

$$\begin{aligned} &\max u, v \left(\frac{u'yi}{v'xi}\right) \\ &s.t. \frac{u'yj}{v'xj} \leq 1 \qquad j = 1, 2, ..., n \\ &u, v \geq 0 \end{aligned} \tag{4}$$

This involves finding values for *u* and *v*, so as the efficiency measure for the *i*-th firm is maximized, subject to the constraints that all efficiency measures must be less than or equal to one. The problem with this ratio formulation is that it has an infinite number of solutions. To avoid this we can impose the constraint $v'xi = 1$, which provides:

$$\begin{aligned} &\max \mu, v\ (\mu'\ yi) \\ &s.t.\ v'\,xi = 1, \quad \mu'\ yi - v \leq 1, \quad j = 1, 2, ..., n \\ &\mu, v \geq 0 \end{aligned} \tag{5}$$

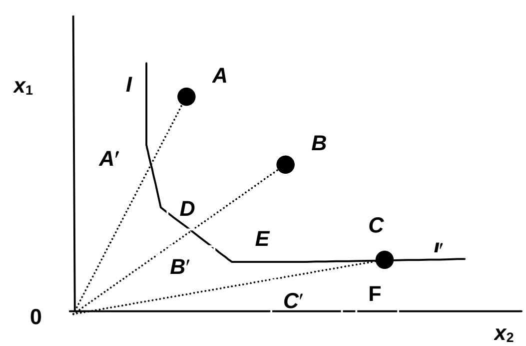
The change of notation of *u* and *v* with μ and ν is used to emphasize that this is a different linear programming problem. This form is called the *multiplier form* of DEA linear programming problem.

Using duality, an equivalent form, the *envelopment form* of this problem can be derived as:

$$\begin{aligned} &\min \theta,\ v\lambda\theta \\ &s.t. - yi + Y\lambda \geq 0 \\ &\theta xi - X\lambda,\ \ v \geq 0 \\ &\lambda \geq 0 \end{aligned} \tag{6}$$

where θ is a scalar and λ is a $Nx1$ vector of constants. The envelopment form has the advantage that it involves fewer constraints and so it is generally preferred. Nevertheless, in several studies, the multiplier form is used, because the weights μ, ν can be interpreted as normalized shadow prices. In the envelop form, θ gives the efficiency score. It takes on values from zero to one. As its value approaches one, firm becomes more technically efficient (observation point is closer to the production frontier). Note that the linear programming problem must be solved *N* times, once for each firm in the sample. A value of θ is then obtained for each firm. The interpretation of the envelopment form problem is as follows. It takes the *i*-th firm and then seeks to radially contract input vector *xi*, as much as possible, while still remaining within the feasible input set. The inner-boundary of this set is a piece-wise linear isoquant, determined by the observed data points. The radial contraction of the input vector, *xi*, produces a projected point ($X\lambda$, $Y\lambda$), on the surface of this technology, which is a linear combination of these observed data points. The constraints in the equation ensure that this projected point cannot lie outside the feasible set. One important problem related to DEA, is slack variables. DEA method, projects the points of inefficient production units to the production frontiers and by doing so, it suggests a combination of inputs that maximize the technical efficiency of the specific firm. The problem of slack variables arises from the fact that a part of the production frontier is parallel to the axis. Because, the DEA method calculates the distance of a firm from the production frontier supposing equiproportional decrease of all inputs, it is possible that a production unit may lie upon the part of the production frontier that is parallel to the axis. In this case, the production unit is technically efficient according to Farrell, but not Pareto efficient. The latter, demands that, keeping output level constant, there is no feasible reduction of any input without the increase of at least one other input. Looking at the following figure, firm *F*, can decrease input $x_2$ keeping output level

**Figure 1**
**Input Efficiency, Radial Measures and Input Slacks**




*Source:* Galanopoulos *et al*., (2005).

constant. Also, firms *A* and *C*, could equiproportionately decrease their inputs until they reach points *A′* and *C′* respectively. But again, it is possible to reduce inputs $x_1$ and $x_2$ respectively, keeping output level constant. So, only the equiproportionate reduction of DMU *B* inputs (reaching point *B′*) is enough to satisfy both the Farrell and Pareto criteria. Equiproportionate reduction of inputs in the case of firms *A, C* and *F* (reaching points *A′*, *C′* and *F′*, respectively) can satisfy only the Farrell criterion. In those cases, slack variables are called input slacks.

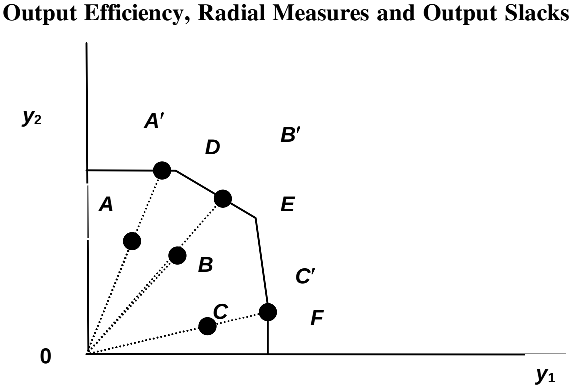
The analysis so far in his paper has been input oriented. However, technical efficiency can be also calculated using output orientation. In this case, it measures the proportion of output level expansion, keeping input level constant. The output-oriented DEA model is similar to the input oriented DEA model. Its envelopment form is as follows:

$$\begin{aligned} &\max \phi\lambda,\ \phi \\ &s.t. -\phi yi + Y\lambda \geq 0 \\ &xi - X\lambda,\ \nu \geq 0,\ \lambda \geq 0 \end{aligned} \tag{7}$$

where $\infty > \varphi \geq 1$ and $\varphi - 1$ is the percentage increase of outputs that production unit *i* can achieve, keeping input levels constant. $1/\varphi$ defines the level of technical efficiency, taking values from 0 to 1. In the CRS case, unlike VRS, as we shall see later, the degree of technical efficiency is the same, independently of the orientation choice. Slack variables may also be present, in the case of output-oriented DEA models. In this case, they are referred as output slacks. As one can see in the following figure, firm *D* and *E* are technically efficient, according to both the Farrell and Pareto criteria, while production unit *F* is technically efficient according to Farrell, but not according to Pareto. That is, it can produce more output $y_2$, keeping input level constant. Production units *A*, *B*, *C,* can equiproportionately increase their outputs, reaching points *A′*, *B′*, *C′*, respectively. For production unit *B*, this increase is enough to satisfy both criteria. This is not the case for production units *A* and *C*, as they can still increase one of their outputs, keeping the other constant.

There are many suggested ways to resolve the problem of variable slacks. Those ways can be distinguished in radial and non-radial measures (Coelli *et al*., 1998). The first category includes the two-stage DEA and the multiple-stage DEA approach, while the second includes the models of Charnes *et al*., (1978) and Färe and Lovell (1978).

**Figure 2**
**Output Efficiency, Radial Measures and Output Slacks**




*Source:* Coelli *et al*., (1998).

Nevertheless, Coelli *et al*. (1998) suggests that the importance of slacks may be overstated. Slacks may be viewed as being an artefact of the frontier construction method chosen and the use of finite sample sizes. If an infinite sample size was available and/or if an alternative frontier construction method was used, which involve a smooth function surface, the slack issue would disappear. Hence, we believe that the analysis of technical efficiency can reasonably concentrate upon the radial Farrell efficiency score.

### 5.2. The Variable Returns to Scale (VRS) DEA Model (Input Oriented VRS DEA)

Imperfect competition, constraints in finance and other reasons, may cause the firm not to operate in the optimal scale. In these cases, CRS assumption is not appropriate resulting in measures of technical efficiency which are confounded by scale efficiencies. For this reason, Banker, Charnes and Cooper (1984), suggests a DEA model which accounts for VRS situation. They introduce an additional convexity constraint: $N_1'\lambda = 1$, where $N1$ is a $Nx1$ vector of ones. This constraint ensures that an inefficient firm is only 'benchmarked' against firms of similar size. That is, the projected point of that firm on the DEA frontier will be a convex combination of observed firms. In CRS DEA models, an inefficient firm may be benchmarked against larger or smaller firms. In this instance the sum of the λ-weights will be greater or smaller than one, respectively. With the addition of the above constraint, the problem takes the form:

$$\begin{aligned} & \min \theta\lambda, \theta \\ & s.t. - yi + Y\lambda \geq 0 \\ & \theta xi - X\lambda \geq 0, \\ & N_1'\lambda = 1 \\ & \lambda \geq 0 \end{aligned} \tag{8}$$

This approach forms a convex hull of intersecting planes which envelope the data points more tightly than the CRS conical hull and thus provides technical efficiency scores which are greater than or equal to those obtained by the CRS model. VRS specification has been the most commonly used specification in the 90's (Coelli, *et al*., 1998).

Like the case of CRS DEA, the VRS DEA model could be output oriented. In this case, the model is as follows:

$$\begin{aligned} & \max \phi\lambda, \phi \\ & s.t. - \phi yi + Y\lambda \geq 0 \\ & xi - X\lambda \geq 0, \\ & N_1'\lambda = 1 \\ & \lambda \geq 0 \end{aligned} \tag{9}$$

where $\infty > \varphi \geq 1$ and $\varphi - 1$ is the percentage increase of outputs that production unit $i$ can achieve, keeping input levels constant. $1/\varphi$ defines the level of technical efficiency, which is between 0 and 1.

Scale efficiency is computed as the ratio of the technical efficiency calculated by the CCR model (assuming constant returns to scale) to the technical efficiency calculated by the BCC model (assuming variable returns to scale) (Banker *et al*., 1984). If this ratio is less than one, the DMU is not producing in the range where CRS holds and potential efficiency gains could be realized with a change of the firm size. However, as noted by Banker and Thrall (1992), the returns-to-scale concept is well defined only for the firms that are technically efficient under the BCC model, but do not lie on the production frontier under the CCR model. For the firms that lie below the production frontier under the BCC model, productivity changes due to returns-to-scale are confounded with productivity changes due to the elimination of inefficiency. Following the above procedure, one can only identify if a firm is scale efficient or not. However, it cannot be concluded that the DMU's scale inefficiency is due to operation in an area of increasing or decreasing returns to scale. This issue can be determined by running an additional DEA model, the Non Increasing Returns to Scale (NIRS) DEA model. This model has the form:

$$\begin{aligned} &\min \theta\lambda, \theta \\ &s.t. - yi + Y\lambda \geq 0 \\ &\theta xi - X\lambda \geq 0, \\ &N_1'\lambda \leq 1 \\ &\lambda \geq 0 \end{aligned} \tag{10}$$

In fact, the form of this model is constructed by the VRS firm will not be benchmarked against larger firms, but it may be benchmarked against smaller firms. When a firm has scale efficiency score equal to one, it is concluded that it operates under constant returns to scale. When a firm's scale efficiency score is less than one, two cases are identified:

1. If the VRS DEA model efficiency score is equal to the NIRS DEA model efficiency score (point *P*), then the scale inefficiency is due to operation under decreasing returns to scale. We can conclude that the firm has to increase its size.
2. If the VRS DEA model efficiency score is higher than the NIRS DEA model efficiency score (point *Q*), the scale inefficiency score is due to operation under increasing returns to scale. We can conclude that the firm has to decrease its size. While the approach for the identification of scale inefficiency is the same in either input- or output-orientation, the results may be different.

While DEA is nonparametric, it is not free of the necessity of modeling the economic theory. Thus, for example, assumptions about the underlying technology will determine whether the efficient frontier is forced through the origin (implying constant returns to scale (CRS)) or allowed not to pass through the origin (implying variable returns to scale (VRS)). The CRS assumption is appropriate only when production is optimal (i.e., corresponding to the flat portion of the long-run average cost curve). A number of factors including, for example, imperfect competition or regulation may cause suboptimal production. The use of the CRS specification when production is not at the optimal level will result in measures of technical efficiency which are confounded by scale efficiencies The use of the VRS specification, as is done here, permits the calculation of technical efficiencies devoid of these scale efficiency effects. To measure both technical and allocative efficiency, it is necessary to have price information on the inputs and be willing to assume that cost minimization is the objective.

In applied research, the choice of input or output orientation has both theoretical and practical implications. Generally, input-orientated DEA models are commonly used. This is because many DMUs have particular orders to fill, so it seems that the input quantities are of main importance. However, a firm's objective may be the maximization of output subject to a fixed level of inputs. In such cases, output-orientated DEA models would be more appropriate. Essentially, one should select the orientation according to which quantities (inputs or outputs) the managers have most control over. An important point to mention is that the output- and input-orientated models will estimate exactly the same frontier and therefore by definition, will identify the same set of firms as efficient. It is only the efficiency measures associated with the inefficient firms that may differ between those two methods.

## 6. DEA: CRITICISMS AND LIMITATIONS

Data Envelopment Analysis (DEA) methodology offers major advantages, since the non-parametric nature of the technique avoids the need to specify beforehand any particular functional form for the technology. Furthermore, this approach does not require any assumption about market structure or about the absence of market imperfections. Furthermore, DEA can deal with the case of multiple input and outputs as well as factors outside the control of individual managements, treating them as fixed inputs (Levitt and Joyce, 1987). DEA places no restrictions on the functional form of the production relationship and makes no a priori distinction between the relative importance of any combination of outputs or inputs. There is also no need to make restrictive assumptions about either the technology representing the production process or the distribution of the component of the residuals which represent inefficiency. Since DEA does not impose any functional form on the data, nor does it make distributional assumptions for the inefficiency term, DEA overcomes some of the specific weaknesses of the other methods, such as a particular functional form for technology, particular assumptions on market structure, and the hypothesis that markets are perfect. The DEA approach also recognizes or can accommodate both discretionary and nondiscretionary inputs and outputs. Technical change as well as network or dynamic assessments can easily be accommodated with DEA. DEA may also accommodate the determination of capacity and capacity output subject to different levels or combinations of multiple outputs (e.g., determination of efficiency, capacity, or capacity utilization when one or more outputs are constrained to zero); it is only necessary to impose constraints in terms of nondiscretionary and weak disposability of inputs or outputs.

Respectively, a recognized limitation of using the DEA to assess technical efficiency is that recommendations for decreasing input usage or expanding output levels are in terms of scalar valued ratios which are held constant (i.e., recommendations are in terms of fixed proportions). This limitation, however, is partially mitigated by considering changes in terms of slack variables. In this case, it is possible to determine decreases in inputs or increases in outputs relative to the slack variables; changes are not restricted to constancy of the input or output mixes. Another option to avoid the problem of constant mix ratios is to consider either an economic cost approach or an economic revenue approach. With the economic DEA approaches, prices on inputs or on outputs are all that are required. Changes to achieve technically and allocatively efficient levels are determined and are not restricted to constant input or output mixes. Further, DEA is a non-parametric approach so does not take into account random error. Hence, it is not subsequently subject to the problems of assuming an underlying distribution about the error term. However, since DEA cannot take account of such statistical noise, the efficiency estimates may be biased if the production process is largely characterized by stochastic elements.

The econometric (SFA) approach estimates a (probably unknown) underlying input–output production relationship using a functional form characterizing the data. Alternative specifications can be statistically tested. On the other hand, DEA programming approach (in imposing no functional form) is less restrictive, but instead takes the bounding observations as defining the best practice efficient frontier. There are, however, no widely adopted tests in mathematical programming to help determine the appropriate selection of inputs and outputs. Any and all variation between observed units and the frontier is attributed to inefficiency while, in the econometric method, that variation is separated into inefficiency and random error. An important advantage of the econometric frontier is that there are a number of well-developed statistical tests to investigate the validity of the model specification – tests of significance for the inclusion or exclusion of factors, or for the functional form. The accuracy of this hypothesis depends to some extent on the assumption of normality of errors, which is not always fulfilled. A second advantage of the econometric frontier is that if a variable which is not relevant is included, it will have a low or even zero weighting in the calculation of the efficient scores, so that its impact is likely to be negligible. This is an important difference from DEA, for which the weight for a variable is usually unconstrained. A third advantage of the econometric frontier is that it allows the decomposition of deviations from efficient levels between 'noise' (or stochastic shocks) and pure inefficiency, while the DEA classifies the whole deviation as inefficiency.

## 7. CONCLUSION

There is an on-going debate among researchers about the applicability and usefulness of the DEA approach vs. the stochastic frontier approach. While usually used to measure technical efficiency (e.g., maximum output from available inputs), both the SFA and the DEA methods can be used to derive allocative efficiency (the least-cost input combination yielding the output) and, thus, overall efficiency measures.

Finally, the efficiency scores from both econometric and programming approaches are often subject to second-stage regression analysis to help determine the impact upon efficiency of environmental factors beyond decision maker control. The success of both approaches relies on some common factors, including that all inputs and outputs are homogeneous across DMUs, are measurable, are measured accurately, are included, and that DMUs are relatively alike and employ a common technology.

Both DEA and SF analysis are popular methods for assessing relative efficiency. Unfortunately, there is no definitive mechanism for selecting between the two. While usually used to measure technical efficiency (e.g., maximum output from available inputs), both the SF and the DEA methods can be used to derive allocative efficiency (the least-cost input combination yielding the output) and, thus, overall efficiency measures. A case can be made for each, and analysts have chosen to use both (although rarely together).

***Acknowledgement***

The author should thank Dr. Richard Harris, Department of Economics, University of Glasgow for all the guidance and contributing remarks. However, any faults are of sole responsibility of the author.

## *Notes*

1. It should be noted, however, that DEA identifies two or more enterprises that represent the best practice of a set of entities. This means that it will always choose a couple or more enterprises as being 100 per cent technically efficient. This could be a limitation in some contexts because it is possible that all enterprises in a sample may be technically inefficient to some extent when compared with a conceptual frontier, and even the best practice enterprises in a sample may still be some distance from being 'fully efficient'. With DEA, the best practice enterprises are defined only relative to other enterprises in the given dataset, and do not necessarily produce output at the potential production frontier.
2. This feature differentiates DEA from the parametric approaches, which require a specific pre-specified functional form of the modelled production or cost function. The Data Envelopment Analysis method builds on the individual firm evaluations, and extends efficiency measures from a single-input, single-output efficiency analysis to multi-input, multi-output situations. (e.g. Charnes, Cooper, Lewin & Seiford 1994, Cooper, Seiford & Tone 2000, Cooper, Seiford & Zhu 2004).
3. Following this early work, many studies have further developed DEA methodology, including those by Charnes *et al*., (1994), and Fare *et al*., (1985, 1994).
4. A cost version of the DEA method refers to the cost or expenditure made to produce the outputs replaces the several physical inputs in the analysis. In this case, *cj*, the cost of DMU *j*, replaces $x_{ij}$ in equation (1). Unpriced multiple outputs remain as in the usual production form. Fare and Grosskopf (1985) show the correspondence between the primal production approach and the dual cost approach. This cost indirect production correspondence is appropriate for evaluating production efficiency when output is not priced but inputs are, so resource usage is reflected in costs. This approach evaluates output provided for given expenditure.

## *References*

Arega D. A., and Manfred Z., (2005), Technology Adoption and Firmer Efficiency in Multiple Crops Production in Eastern Ethiopia: A Comparison of Parametric and Non-Parametric Distance Functions, *Agricultural Economics Review*, 6(1), 5-17.

Banker R. D., and Thrall R. M., (1992), Estimation of Returns to Scale Using Data Envelopment Analysis, *European Journal of Operational Research*, 62, 74-84.

Banker R. D., Charnes A., Cooper W. W., and Maindiratta A., (1988), A Comparison of DEA and Translog Estimates of Production Frontiers Using Simulated Observations from a Known Technology, In: Dogramaci A., & Fare R., (Eds), *Applications of Modern Production Theory: Efficiency and Productivity*, Boston, MA, Kluwer.

Charnes A., Cooper W. W., Lewin A. Y., and Seiford L. M., (1994), *Data Envelopment Analysis: Theory, Methodology and Application*, Kluwer Academic, London.

Charnes A., Cooper W. W., and Rhodes E., (1978), Measuring the Efficiency of Decision Making Units, *European Journal of Operational Research*, 2, 429-444.

Charnes A., Cooper W. W., Lewin A. Y., and Seiford L. M., (1993), *Data Envelopment Analysis: Theory, Methodology and Applications*, Boston, MA, Kluwer.

Coelli T., Rao D. S. P., and Battese G., (1998), *An Introduction to Efficiency and Productivity Analysis*, Boston, MA, Kluwer.

Coelli T. J., (1995), Recent Developments in Frontier Modelling and Efficiency Measurement, *Australian Journal of Agricultural Economics*, 39(3), 219-245.

Cooper W. W., Seiford L. M., and Zhu J., (2004), Data Envelopement Analysis: Models and Interpretations, Chapter 1, 1-39, In W. W. Cooper, L. M. Seiford, and J. Zhu, (Eds), *Handbook on Data Envelopment Analysis*, Kluwer Academic Publisher, Boston.

Cooper W. W., Seiford L. M., and Tone K., (2000), *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, Kluwer Academic Publishers, Boston.

Desai A., and Schinnar A. P., (1987), Technical Issues in Measuring Scholastic Improvement Due to Compensatory Education Programs, *Socio-Economic Planning Sciences*, 24(2), 143-153.

Fare R., Grosskopf S., and Lovell C. A. K., (1985), The Measurement of Efficiency of Production, Boston, Kluwer-Nijhoff.

Fare R., and Lovell C. A. K., (1978), Measuring the Technical Efficiency of Production, *Journal of Economic Theory*, 19, 150-162.

Färe Rolf, Shawna Grosskopf, and C. A. Knox Lovell, (1994), Production Frontiers, Cambridge, Cambridge University Press.

Färe Rolf, Shawna Grosskopf, Mary Norris, and Zhongyang Zhang, (1994), Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries, *American Economic Review*, 84, 66-83.

Farrell M., (1957), The Measurement of Productive Efficiency, *Journal of the Royal Statistical Society Series A (General)*, 120 (3), 253-281.

Galanopoulos K., Aggelopoulos S., Kamenidou I., and Mattas K., (2005), Assessing the Effects of Managerial and Production Practices on the Efficiency of Commercial Pig Firming, *Agricultural Systems*, (Accepted for Publication).

Hjalmarsson L., Kumbhakar S. C., and Heshmati A., (1996), DEA, DFA and SFA: A Comparison, *Journal of Productivity Analysis*, 7(2-3), 303-327.

Iráizoz B., Rapún M., Zabaleta I., (2003), Assessing the Technical Efficiency of Horticultural Production in Navarra, Spain, *Agricultural Systems*, 78, 387-403.

Korres G., (2007a), *Technical Change and Economic Growth: Inside to the Knowledge Based Economy: Recent Evidence on European Perspectives*, Avebury-Ashgate, London.

Korres G., (2007b), Industrial and Innovation Policy in Europe: The Effects on Growth and Sustainability, *Bulletin of Science, Technology and Society*.

Levitt M. S., and Joyce M. A. S., (1987), The Growth and Efficiency of Public Spending, The National Institute of Economic and Social Research, *Occasional Papers XLI*, Cambridge University Press, Cambridge.

McMillan M. L., and Chan W. H., (2006), University Efficiency: A Comparison and Consolidation of Results from Stochastic and Non-Stochastic Methods, *Education Economics*, 14(1), 1-30.

Serrao A., (2003), *Agricultural Productivity Analysis of European Union and Eastern Regions, Paper Prepared for Presentation at the American Agricultural Economics Association (AAEA) Annual Meeting*, Montreal, Canada, July 27-30, 3-20.

Sengupta J. K., (1987), Data Envelopment Analysis for Efficiency Measurement in the Stochastic Case, *Computers and Operations Research*, 14(2), 117-129.

Sharma K. R., Leunga P., and Zaleski H. M., (1999), Technical, Allocative and Economic Efficiencies in Swine Production in Hawaii: A Comparison of Parametric and Nonparametric Approaches, *Agricultural Economics*, 20, 23-35.