ASSESSING PUBLIC SECTOR EFFICIENCY: ISSUES AND METHODOLOGIES

Francisco Pedraja-Chaparro[*]*, Javier Salinas-Jiménez*[**] *and Peter C. Smith*[***]

Introduction

A more efficient public sector has become a universal target of central importance in economic policy. The increase in both the quantitative and qualitative relevance of the public sector within the economy (in terms of its size and functions), its contribution of budgetary discipline in the pursuit of macroeconomic stability and the difficulty of increasing public revenues are some of the reasons why attention is focusing on the public expenditure side of Public Finance.

A whole variety of initiatives ranging from privatisations to market simulations have been implemented in the pursuit of public sector efficiency. The introduction into the public sector of private sector management techniques such as decentralisation, management and performance measures, customer services, and so on comprise what is known as the New Public Management. These initiatives in the reform process have been implemented with varying levels of scope and intensity in western countries. The pioneers have been the UK and the USA, and the most demanding and comprehensive case has been New Zealand. The success of these initiatives to improve efficiency in the public sector depends crucially on the extent and confidence with which we are able to measure the performance of public services.

A central concern is to measure the relative efficiency of different public organisations providing the same public service. Two alternative approaches can be considered. The first is to develop a set of performance indicators, *i.e.* partial measures of some aspects in the behaviour of the organisation. The second is to try to develop a general index on the efficiency of the organisation. Although the first alternative has some virtues, its main flaws lie in its partial nature. As a consequence, contradictory results may arise, depending on the choice of indicator (Smith, 1990 and Smith and Goddard, 2003). The development of global efficiency scores seeks to overcome this weakness.

The traditional productivity literature characterises global measures of organisational efficiency as the distance of the unit under scrutiny from a frontier function, which is estimated using the best observed practice of the set of other similar units. There are two main methodologies for defining the frontier (Green and Coelli, 1998).

[*] Universidad de Extremadura (Spain).

[**] Instituto de Estudios Fiscales – Universidad de Extremadura (Spain).

[***] Centre for Health Economics – University of York (United Kingdom).

Parametric approaches that specify *a priori* a functional form with constant parameters to be estimated (*i.e*. Cobb-Douglas, translog, etc.). Efficiency is assessed in relation to this function and will be different depending on the chosen functional form. Stochastic-type frontiers are generally used where two components are identified in the residuals: inefficiency and all the other sources of error.

Non-parametric approaches that do not *a priori* specify a functional form, but nevertheless require certain assumptions about the structure of the production technology (e.g. free disposability, convexity, etc). This is done by solving a separate mathematical programming model for each observation. As in the parametric approach, the frontier changes, as will the estimated efficiency of each unit, in line with the assumptions made. This kind of approach generally is of a deterministic type, with all the distance from the frontier assumed to be caused by inefficiency.

Data Envelopment Analysis (DEA) is the dominant non-parametric technique in productivity analysis. Since it was introduced in 1978, it has undergone substantial theoretical development, and enjoyed a rapid growth in empirical applications in diverse fields, amongst which the public sector is very important.

The increased availability of data related to public sector performance and the ready availability of software enabling us to implement the DEA model raises the question of how to interpret efficiency scores obtained for different units for the purposes of improving their future performance. This paper therefore seeks to assess the usefulness of DEA in measuring the efficiency of a set of comparable units in the public sector. We must also note that our objective focuses exclusively on the extent to which DEA can yield an adequate measure of efficiency, and not on the implications of using such measures in a targets or reward scheme.

The paper is structured as follows. Some reflections are made on the special characteristics of public units' performance and the way in which these characteristics might influence the assessment of their efficiency. Secondly, the DEA model is briefly described. Finally, we offer some reflections on methodological issues that seem significant when assessing the efficiency of a set of public units through such non-parametric approaches.

1. Characteristics of "public supply" and efficiency

It is usual to start any discussion of public services by emphasising their unique characteristics, most especially the absence (or near absence) of any market in the conventional sense. When calculating the efficiency of a set of public units, we must therefore first briefly consider the special characteristics of public sector supply. This will shed light on some of the measurement and conceptual issues we are faced with.

A first restriction arises from the nature of the objective function for the public sector which is characterised by multiple criteria. In addition to efficiency, public sector activities often try to achieve equity goals, and there often exists a trade off among these objectives. Diverse and conflicting objectives are components that must be borne in mind in any assessment of a public policy, to avoid hasty conclusions made when scrutinising only one of them. In addition to the existence of multiple objectives, the public sector differs from the private sector because of the diversity of principals (politicians, users, general public) that must be satisfied by agents (or 'bureaucrats', to use the terminology of the public choice literature). The multiplicity of tasks and principals causes serious problems when measuring public output.

Public output cannot be traded in the market, so it is difficult to define and to measure it. In practice, most of the time we resort to metrics in which aspects that are difficult to calculate (such as those related to the quality of the service) are not considered. The lack of a market prevents consumers from expressing a valuation of the services. Instead, the value of public output must often be inferred by observing public service activities. At best we can usually deal only with intermediate outputs (often with variables closer to measuring inputs than outputs) or mere proxies for final outputs (the services' effects on users). In addition to the measurement problems, there exist problems with attribution, in the sense that outputs may result from factors totally or partially out of the control of assessed units. In any case, the public sector is not a uniform body. Services and organisations of a different nature coexist, from the simplest ones to the most complex, where measurement and allocation problems appear at a different level.

The measurement and attribution issues lead to a monitoring and control problem. How should production be regulated? The regulatory problem is one of inferring optimal production in the absence of competitive pressure. In the absence of good regulation, public service units are likely to exhibit inefficient behaviour, in both an allocative and productive sense. However, the lack of competition can make the production technology uncertain and unknown. The question therefore is: does a comparison base exist that can furnish information on the technology?

Finally, there is in the public services an absence of the entry and exit options manifest in competitive markets. In particular, there is no guarantee that inefficient producers of public output are subject to the threat of bankruptcy that acts as a discipline in competitive markets.

The absence of competition, the monopolistic nature of public production and the absence of bankruptcy threat are some explanations for the difficulties in regulating public production. Moreover, or rather consequently, the schemes of internal incentives (positive or negative) found in most public services cannot guarantee efficient production. In the light of the preceding discussion, any measurement technique should be adapted to the following characteristics:

The lack of a market and the resulting difficulty of measuring the actual output make us use an intermediate output. Consequently, the technique must accommodate a measurement problem characterised by multiple output and input.

The technique should adjust to the characteristics of uncertainty surrounding public production technologies. Thus, it is advisable to use approaches which are flexible and do not require very strong assumptions on the production frontier.

The purpose of this paper is to assess the extent to which data envelopment analysis can assist regulators confronted by these difficulties, and seeking to assess the performance of public service institutions.

2. Data envelopment analysis (DEA)

DEA is a mathematical programming procedure developed by Charnes, Cooper and Rhodes (1978), based on the seminal work by Farrell (1957). The DEA model applies mathematical programming techniques to compare the efficiency of a set of units. DEA may be seen as an extension of the traditional output/input ratio analysis. The efficiency score of each unit can be represented as a ratio of the total weighted outputs to the total weighted inputs:

$$\text{Efficiency} = \frac{\text{Weighted sum of outputs}}{\text{Weighted sum of inputs}}$$

Mathematically, if we consider a set of n units consuming m inputs ($x_1 \ldots . x_m$) and producing s outputs ($y_1 \ldots . y_s$), the efficiency of a unit, say unit 0, can be measured as follows:

$$\text{Maximize } h_0 = \frac{\sum_{r=1}^{s} U_{ro} Y_{rj}}{\sum_{r=1}^{m} V_{io} X_{ij}} \leq 1$$

$$\text{subject to: } \frac{\sum_{r=1}^{s} U_{ro} Y_{ro}}{\sum_{r=1}^{m} V_{io} X_{io}} \qquad j = 0, 1, 2, \ldots, n$$

$$u_{ro} \geq 0; \qquad r = 1, 2, ..., s$$

$$v_{io} \geq 0; \qquad i = 1, 2, ..., m$$

By solving this problem, it is possible to calculate, for each one of the units, the set of inputs and outputs weights with which the unit may obtain the maximum efficiency score, with the restriction that using the same weights no other unit can achieve an efficiency score higher than one.

If, subject to this restriction, it is possible to find a set of weightings in which the efficiency ratio of the unit analysed equals one, that unit will be considered as efficient. Otherwise, the unit will be evaluated as inefficient since, even if a more favourable set of weights is considered, one could find another unit obtaining a greater efficiency ratio. In assessing efficiency in this way, weightings assigned to inputs and outputs will vary by unit. As pointed out by Sexton (1986), since each unit uses differing combinations of inputs and outputs, a different set of weightings will in general be selected for each – that is, the set allowing the unit to obtain the greatest efficiency ratio. Therefore, the method assesses each unit in the most favourable light.

Presenting DEA as a weighted sum of outputs in relation to a weighted sum of inputs allows one to characterize the technique as an extension of the analysis of ratios and, as pointed out in the introduction, to approach efficiency from a global point of view.

The maximization problem specified above can be presented in a linear form, which is more convenient for solving, as follows:

$$\text{Maximize } \sum_{r=1}^{s} u_r y_{r0}$$

$$\text{subject to: } \sum_{i=1}^{m} v_i x_{i0} = 1$$

$$\sum_{r=1}^{s} \mu_r \gamma_{rj} - \sum_{i=l}^{m} v_i x_{ij} \leq 0 \qquad j = 1, 2, \ldots, n$$

$$u_{ro} \geq 0; \qquad r = 1, 2, ..., s$$

$$v_{io} \geq 0; \qquad i = 1, 2, ..., m$$

The dual problem is:

$$\text{Minimize } \theta_0$$

subject to:

$$\sum_{j=1}^{n} x_{ij} . \lambda_j - x_{i0}\theta_0 \leq 0 \qquad i = 1, 2, \ldots, m$$

$$\sum_{j=1}^{n} y_{rj} . \lambda_j \geq y_{r0} \qquad r = 1, 2, \ldots, s$$

with trivial restrictions $\theta_0, \lambda_j \geq 0$ $\qquad j = 1, 2, ..., n$

The interpretation of this programme is very simple, as with the fractional programme. The aim is to search, for each assessed unit, a linear combination of other units that produces at least the same quantity of output in each of the s dimensions considered by consuming a lesser proportion ($0<\theta<1$) of the m inputs. If this is not possible, the unit is efficient. On the other hand, if it is possible, the unit will be assessed as inefficient since there are other units in the sample (those from the reference group against which it is compared) performing better.

3. Assessment of efficiency in the public sector: some methodological reflections

Because *a priori* the production technology is unknown and therefore there is no single optimal approach towards assessing the efficiency of a set of productive units, we hesitate to advocate one particular analytic technique over another. In general, the characteristics of the sector analysed, together with the restrictions of information and the purpose of the analysis will determine the most suitable technique. Nevertheless, DEA seems suitable for the multidimensional character of public output, and its flexibility is particularly attractive given the lack of knowledge and the uncertainty involved in the public sector production process.

In this section we therefore discuss some of the most significant methodological issues in public service efficiency measurement that arise when using a non-parametric approach. In particular, we will analyse issues derived from:

1) the fact that the units may have different objectives;
2) from exogenous and environmental factors (out of the control of units) which may considerably influence their results; and
3) from limitations that arise from the deterministic and non-parametric nature of the DEA model.

3.1 Diversity of objectives and weights

One of the main strengths of DEA is that it enables us to analyse the global performance of the assessed units. In order to assess the efficiency of the units, we have to assume that they are homogenous, consuming the same inputs, producing the same outputs and trying to achieve similar objectives. However, in general, the objectives and outputs are not well defined and, even when they are, it is difficult to quantify them in practice. The analyst often has to turn to the only available data that approximately inform her on the performance of units. In this sense, let us think of the difficulties in areas such as education, health care or justice where final outputs might be the value added in schools, the health gains in health care or the protection granted by courts. The immense difficulties involved in directly or indirectly measuring these outcomes usually leads to the use of proxies such as pupils' results, the number of patients or the number of resolved cases respectively.

The case of decentralised services is particularly complex because we cannot guarantee uniformity in the objectives of the units under scrutiny. Indeed, some of the major gains in efficiency arising from decentralisation are said to derive from a better adjustment to differentiated preferences for these public services in the different jurisdictions, which contradicts the above identity of objectives assumed in DEA.

However, although the objectives are assumed to be universal in DEA, the weights attached to each input and output are calculated so that each assessed unit receives the most favourable possible treatment. This implies that there are no *a* *priori* values or restrictions set on the different weights and that weights attached to the different inputs and outputs may differ when assessing the various units. Yet, although consistent with notions of decentralisation, total weight flexibility, considered as a significant advantage of DEA (Cooper *et al.*, 2000), may be criticised for various reasons:

- if the inputs and outputs included in the analysis are not equally important it is not sensible to claim that a unit is relatively efficient when the weights to the important inputs and outputs are zero. The total flexibility of the DEA model may lead to an unfounded emphasis on the efficient use of relatively unimportant factors, concealing inefficiencies in the most important activities carried out by the units;
- if we use the unbounded model, some inputs and outputs may be ignored in the analysis when assessing the relative efficiency of some units. As a result, the relative efficiency of a unit may not really reflect its performance in relation to inputs and outputs taken as a whole;
- the implicit assumption made when allowing weight flexibility in the unbounded DEA model is that the analyzed units may have individual objectives and particular circumstances that should be considered when assessing them. However, since the units compared using DEA are homogeneous units, it may be unacceptable to assume that the relative importance attached to the different inputs and outputs by each unit should differ greatly.

Therefore, there seems to be a strong case for considering the introduction of weight restrictions. The research has focused on technical aspects, setting the limits between which the weights can vary and evaluating how the introduction of weight restrictions may improve the results of the DEA model (See Dyson and Thanassoulis, 1988 and Pedraja-Chaparro *et al.*, 1997).

3.2 Exogenous factors

When the efficiency of a set of public units is assessed, the conventional DEA model implies that there is no factor outside the control of the units in charge of providing the public service. This assumption is quite often very far from what actually happens in most public services.

We must distinguish between environmental and exogenous factors among those ones which are outside the control of productive units.

Environmental factors are not directly involved in the production process although they may provide useful information on how to explain different efficient behaviours. These are, among others, the kind of ownership, the degree of competition, geographical factors, etc.

We must particularly deal with exogenous factors that affect the production process but not entirely under the control of units. It would not make any sense to introduce an objective for some unit (*i.e.* a percentage decrease in inputs consumption for it to be efficient), if these inputs were beyond the control of the unit. This type of factor is often found in a number of public services as, for instance, in educational services when pupils' skill and social economic background are more important determinants of results than the resources consumed by schools. DEA has shown a noteworthy adaptability to include these exogenous factors in efficiency assessment.

There are two broad approaches to including factors beyond the control of units. On the one hand, we have the one-stage approach, where the exogenous factors are included jointly with all the others inputs that can be controlled by the productive units. In this way, only one DEA analysis is run, in which all the inputs are included together. The principal model following this approach is that proposed by Banker and Morey in 1986. Its main advantage is that it enables us to introduce all relevant variables in a DEA single analysis, which simplifies to a large extent the calculation of efficiency indices. However this methodology has shortcomings. Units operating in the most disadvantaged circumstances will automatically be deemed efficient regardless of their performance (because there are no direct comparators). More generally, the increased number of variables introduced into the DEA model reduces its power to discriminate between units.

The second alternative is a multi-stage analysis. These models consist of several analytic stages. All of them have in common a first stage in which we only include those inputs that units can control. Afterwards, some adjustments on initial efficiency scores are made, avoiding biases that would lead to benefit unfairly units working in a relatively more favourable context. We can use different methodological options:

- *Two-stage models*. In these models, efficiency scores calculated in the first stage are included as dependent variables in a regression where explanatory variables are non-discretionary inputs. Although there are many ways to undertake this regression, the methodology proposed by Ray in his study of Connecticut schools in 1991 may be highlighted. Following this methodology corrected ordinary least squares are used to obtain consistent estimators of parameters. Their major appeal lies in that this correction guarantees that the units with the worst supply of non-controllable inputs enjoy the largest upwards adjustments.
- *Three-stage model* (Fried and Lovell, 1996). In this model, total slacks obtained in the first stage are included in a second DEA as controllable inputs, whereas outputs are non-controllable inputs. The aim of this second analysis is to identify the part of slacks that can be explained by exogenous factors and the part which reflects technical inefficiency. After separating both influences, the initial values of inputs and outputs are adjusted and then a third DEA model is run using these adjusted values.
- *Four-stage model* (Fried, Schimidt and Yaisawarng, 1999). This methodology can be considered as a mix between the three-stage model (since it also uses total slacks) and the two-stage approach (because slacks are included as dependent variables in a regression with non-controllable factors as explanatory variables). However, in this case a Tobit regression is used instead of ordinary least squares and only inputs slacks are included in regressions, one for each variable.

These models represent rational extensions to the basic DEA model. However, they too can be criticised. In our opinion, an important restriction comes from the possible bias in the results when there is a correlation between the inputs included in the fist stage and the independent variables considered in the second (see Chalos, 1997).

Moreover, two-stage and four-stage models have notable disadvantages, put forward by Simar and Wilson (2003). Specifically, there are problems related to the fact that DEA efficiency estimates are dependent in the statistical sense (they are computed using linear programming techniques) and, consequently, standard approaches to inference are invalid. They suggest employing bootstrap methods in order to overcome these problems.

Thus, in spite of the versatility and adaptability of DEA in handling non-discretionary inputs, there is no generally accepted methodology as to the appropriate way to introduce them when measuring efficiency. The analyst must often use judgement in the light of the characteristics of the specific area of application, data availability and a search for simplicity. This is especially relevant if very different results are obtained when applying the alternative approaches to the same sample. (See Cordero, Pedraja-Chaparro and Salinas-Jiménez, 2004).

3.3 Problems derived from the non-parametric and deterministic nature of DEA

One of the most serious shortcomings of DEA arises from the non-parametric and deterministic nature of the model. In this sense, the following issues are especially relevant:

i) the sensitivity of the results to model specification;

ii) the use of inappropriate data;

iii) the fact that efficiency estimates are point estimations; and

iv) the lack of adequate techniques for treating missing data.

a) Selection of variables and the specification of the model

The analyst faces two fundamental choices when assessing a set of units using DEA: on the one hand, the selection of variables that must be included in the efficiency analysis; and on the other, the type of returns to scale (constant or variable ones) that must be considered in the production function.

Given the deterministic and non-parametric nature of DEA, the choice of variables is a crucial decision that may considerably affect the results obtained in the analysis. As opposed to econometric models, where the analyst can use tests such as the t-test, one cannot apply any model selection test to DEA, and the researcher does not know if the results are robust or if they exclusively arise from the choice of variables used in the analysis.

In this sense, some studies have focused on the comparison of the results of a set of alternative models in order to prove the "validity" of the efficiency estimations. Thus, Gong and Sickles (1992) and Banker, Gadh and Gorr (1993), among others, compare efficiency indices obtained by DEA with those resulting from the application of alternative approaches. Nevertheless, this validation process of results has severe limitations. This is due to the fact that the results that are compared are derived from approaches based on very different assumptions related to the frontier production. Other authors have made a sensitivity analysis of results by calculating efficiency indices with several sets of variables and specifications. This is the path followed by Färe, Grosskopf and Weber (1989) when analysing educational centres and by Valdmanis (1992) when assessing a set of hospitals. However, if the results obtained were sensitive to the specification of the model it would not yet be clear what should be done, apart from relying on the analyst's common sense, bearing in mind the non-parametric nature of technique.

With the same aim, but on the theoretical level, Smith (1997) uses diverse models with simulated data in order to analyse the effects derived from model misspecification in DEA. The major conclusion is that errors derived from model misspecification are larger when the model is simple (*i.e.* with a small number of variables) and the sample is small; in such circumstances, it would be better to include non-relevant variables than leave major variables out of the model.

As regards to the type of returns to scale, it must be pointed out that in order to ensure the homogeneity of units studied in comparisons, the DEA model enables us to specify the type of returns to scale, including this assumption on the building of an efficient frontier. This aspect turns out to be crucial because when it is not considered we would mistake some inefficiency for scale problems. For instance, a wrong use of an assumption of variable returns to scale may favour units that operate at unusually large or small scale, making them incorrectly assessed as efficient. Thus, the first issue the analyst must face, in this sense, is to determine the type of returns to scale in order to estimate the production frontier. Previous studies on the sector and its characteristics may provide a first hint of the type of returns to scale that must be considered. Also, there are several alternative methods in the empirical literature that aim at contrasting the validity of the assumption on returns to scale. Among these alternatives are the following:

- to assume constant returns to scale and analyse the relation between efficiency indices and the size of the assessed units. For this, usually the Tobit model is used because the values of the dependent variable (efficiency indices) are included in the interval 0-1.
- to compare the similarity between results obtained under the assumption of constant returns to scale and that of variable returns, thus calculating in this way possible scale inefficiencies and in consequence finding the most convenient type of assumption for the analysis.

In those cases, where the production function is better known and the production process is simpler, one may use (as a complement) other types of parametric approaches in order to contrast the kind of returns to scale.

b) Inappropriate data

With regard to the use of inappropriate data, major problems arising in empirical papers result from measurement errors or outliers that may distort the efficiency analysis and from the relative small number of observations.

The problem of measurement errors may be abated, provided that they are occasional and are not repeated in the same units reiteratively, making the efficiency analysis multi-period; *i.e.* by repeating the efficiency analysis for different periods of time with one of the following tools:

- analysis of several periods of time and presentation of average results;
- calculation of average values for each variable in some period of time and subsequent efficiency assessment;
- use of the window analysis, a more sophisticated approach that involves considering observations of a same unit in different periods of time as if they were distinct units one from another.

In the public sector, a multi-period analysis is desirable not only because of the non-stochastic nature of the DEA model, but also because of the nature of the expenditure programmes surveyed. When analysing sectors such as education and health services, where the resources used may have medium and long-term effects, it would be advisable that the efficiency analysis were referred to relatively long periods of time.

Naturally, the data with greater impact on the efficiency analysis is that related to units on the frontier (efficient units) as they may affect the assessment of some inefficient units in the sample. For this reason, a range of more or less elaborated methods have been designed to detect which units, among those assessed as efficient, are especially influential. Among these methods we may highlight:

- superefficiency indices, built by Andersen and Petersen (1993) and Wilson (1995) which indicate to which extent an efficient unit is far from the frontier made up by the other units. A very high super efficiency index may be *prima facie* a cause for questioning the efficiency of this unit and for carrying out a particular scrutiny of its data.
- calculation of the number of times an efficient unit appears in reference groups of inefficient units; this device is a basic way to assess the influence of a unit on the others, paying special attention to measurement errors made by the most influential units.

As regard to the relatively small number of units in relation to the number of variables included in the efficiency analysis, there may be a very significant reliability problem referred to the results obtained through the DEA model. The number of free dimensions decreases as new variables are included and, consequently, it is more likely that each unit may be considered efficient because of the model flexibility. In spite of the significance of these aspects, only a small number of papers have dealt with issues related to the selection process of variables and the analysis of degrees of freedom which must exist so that the results in the study may be considered reliable.

Unfortunately, in most empirical studies the choice of variables usually is too dependant on available data; and subsequent haste to apply the technique obviates basic theoretical aspects that should be considered. The criterion suggested by Banker *et al.* (1989) is usually used. They point out, as a rule of thumb, that the number of assessed units should be at least three times the number of variables included in the model. This rule, although it is an *ad hoc* criterion without any theoretical or empirical basis, has been used in many applied studies and considered as a valid criterion to ensure the reliability of results obtained, irrespective of the objectives of the efficiency analysis. Pedraja-Chaparro *et al*. (1999) show how the reliability of results obtained with the DEA model depends not only on the number of observations and variables but also on the distribution of actual efficiencies and on the correlation between inputs and outputs. Generally, Banker's rule, which considers only two of these four aspects, seems too generous. Nevertheless, it may be useful if the objective of the efficiency analysis is to obtain global information from one sector (average efficiency). But Banker's rule may lead to making mistakes if more detailed information is needed (such as efficiency rankings or individual efficiency indices of the different units). Some authors, such as Adler and Golany (2001), have decided to use principal components as outputs and inputs, making it possible to include information from a large set of variables but reducing problems related to the lack of degrees of freedom.

c) Point estimations

Another significant restriction of the DEA model, derived from its non-parametric nature is that the technique only enables us to obtain point efficiency indices of the units. So it is not possible to analyse if the differences between two estimates are statistically significant or make inferences. Wilson and Simar, Löthgren and Tambour (1997) and Simar and Wilson have proposed in recent years the use of "bootstrapping" techniques in order to overcome this restriction and build confidence intervals for the efficiency indices in order to make more accurate comparisons between the assessed units. In the same sense, Ferrier and Hirschberg (1997 and 1999) have proved that bootstrapping techniques may be applied to DEA efficiency scores. Various empirical papers dealing with efficiency measurement in the public sector have used this approach, among which the efficiency assessment in British primary care centres (Giuffrida, 1999) or the study on community care in England (Salinas-Jiménez, Pedraja-Chaparro and Smith, 2003) may be highlighted.

d) Missing data

Finally, due to the nonparametric nature of the model, missing data are a significant problem in DEA. In many empirical applications, blank entries for the data matrices are directly eliminated before the efficiency analysis, However, units with missing data could be highly useful as reference or benchmark units, which span the efficient frontier and eliminating them may distort the efficiency evaluation. Only a few attempts to solve this problem can be found in the DEA literature. Among them, the paper by Kuosmanen (2002) uses dummy variables (zero for missing outputs; number large enough for inputs) and runs a DEA model with weight restrictions in such a way that the black entries are not considered. An alternative approach is proposed by Kao and Liu (2000) who use fuzzy sets to model the ranges for missing data.

4. Conclusions

The theoretical underpinnings DEA have developed to an extraordinary extent since the publication of the initial 1979 paper. However, there remain unresolved many conceptual and operational difficulties associated with the technique, as summarised in this paper. We nevertheless believe that DEA can furnish regulators with useful insights into the performance of public service organisations, if used with discretion, and viewed in conjunction with other techniques. It is especially useful for exploring complex datasets and identifying exceptionally good or poor performers. In its current state of development, it is not suitable for making definitive judgements on organisations or setting detailed targets.

REFERENCES

Adler, N. and B. Golany (2001), "Evaluation of Deregulated Airline Networks Using Data Envelopment Analysis Combined with Principal Components Analysis with an Application to Western Europe", *European Journal of Operational Research,* No. 132, pp. 260-73.

Andersen, P. and N.C. Petersen (1993), "A Procedure for Ranking Efficient Units in Data Envelopment Analysis", *Management Science*, Vol. 30, No. 10, pp. 1261-64.

Banker, R.D., A. Charnes, W.W. Cooper, J. Swarts and D.A. Thomas (1989), "An Introduction to Data Envelopment Analysis with Some of Their Models and its Uses", *Research in Governmental and Nonprofit Accounting*, No. 5, pp. 125-63.

Banker, R.D., V.M. Gadh and W.L. Gorr (1993), "A Monte Carlo Comparison of Two Production Frontier Estimation Methods: Corrected Ordinary Least Squares and Data Envelopment Analysis", *European Journal of Operational Research,* No. 67, pp. 332-43.

Banker, R.D. and R.C. Morey (1986), "Efficiency Analysis for Exogenously Fixed Inputs and Outputs", *Operations Research,* Vol. 34, No. 4.

Bosch, N., F. Pedraja-Chaparro and J. Suárez Pandiello (2000), "Measuring the Efficiency of Spanish Municipal Refuse Collection Service", *Local Government Studies,* Vol. 26, No. 3, pp. 71-90.

Coelli, T., D. Rao and G. Battese (1998), "An Introduction to Efficiency and Productivity Analysis", Boston, Kluwer Academic Publishers.

Cordero, J.M., F. Pedraja-Chaparro and J. Salinas-Jiménez (2004), "Efficiency Measurement in Education and Non-discretionary Inputs", presented at the *North American Productivity Workshop 2004*, Toronto, Canada, June.

Chalos, P. (1997), "An Examination of Budgetary Inefficiency in Education Using Data Envelopment Analysis", *Financial and Accountability & Management*, No. 13, pp. 55-69.

Charnes, A., W.W. Cooper and E. Rhodes (1978), "Measuring the Efficiency of Decision Making Units", *European Journal of Operational Research,* Vol. 2, pp. 429-44.

Cooper, W.W., L.M. Seiford and K. Tone (2000), "Data Envelopment Analysis: A Comprehensive Text With Models, Applications and References", Boston, Kluwer.

Cullis, J.G. and P.R. Jones (1987), "Microeconomics and the Public Economy: A Defence of Leviathan", Blackwell.

Dixit, A. (2002), "Incentives and Organizations in the Public Sector: An Interpretive Review", *Journal of Human Resources,* Vol. 37, No. 4, pp. 696-727.

Downs, A. (1957), "An Economic Theory of Democracy", New York, Harper and Row.

Dyson, R. and E. Thanassoulis (1988), "Reducing Weight Flexibility in Data Envelopment Analysis", *Journal of the Operational Research Society*, Vol. 39, No. 6, pp. 563-76.

Färe, R., S. Grosskopf and W.L. Weber (1989), "Measuring School District Performance", *Public Finance Quartely,* Vol. 17, No. 4, pp. 409-28.

Farrell, M.J. (1957), "The Measurement of Productive Efficiency", *Journal of the Royal Statistical Society* (A), Vol. 3, pp. 253-90.

Ferrier, G.D. and J.G. Hirschberg (1997), "Bootstrapping Confidence Intervals for Linear Programming Efficiency Scores: With an Illustration Using Italian Banking Data", *Journal of Productivity Analysis,* Vol. 8, pp. 19-33.

Ferrier, G.D. and J.G. Hirschberg (1999), "Can We Bootstrap DEA Scores?", *Journal of Productivity Analysis,* Vol. 11, pp. 81-92.

Frantz, R. (1988), "X-Efficiency: Theory, Evidence and Applications", Boston, Kluwer.

Fried, H.O. and C.A.K. Lovell (1996), "Searching for the Zeds", Georgia Productivity Workshop.

Fried, H.O., S. Schmidt and S. Yaisawarng (1999), "Incorporating the Operating Environment into a Non-parametric Measure of Technical Efficiency", *Journal of Productivity Analysis*, No. 12, pp. 249-67.

Fried, H.O., C.A.K. Lovell, S.S. Schmidt and S. Yaisawarng (2002), "Accounting for Environmental Effects and Statistical Noise in Data Envelopment Analysis", *Journal of Productivity Analysis*, Vol. 17, pp. 157-74.

Green, W.H. (1993), "The Econometric Approach to Efficiency Analysis", in H.O. Fried, C.A.K. Lovell and S.S. Schmidt (eds.), *The Measurement of Productive Efficiency: Techniques and Applications*, New York, Oxford University Press.

Gong, B.H. and R.C. Sickles (1992), "Finite Sample Evidence on the Performance of Stochastic Frontiers and Data Envelopment Analysis Using Panel Data", *Journal of Econometrics,* Vol. 51, pp. 259-84.

Kao, C. and S.T. Liu (2000), "Data Envelopment Analysis with Missing Data: An Application to University Libraries in Taiwan", *Journal of the Operational Research Society,* Vol. 51, No. 8, pp. 897-905.

Koopmans, T.C. (1957), Three Essays on the State of Economic Science, New York, McGraw Hill.

Kuosmanen, H. (2002), "Modelling Blank Data Entries in Data Envelopment Analysis", Working Paper, Department of Social Sciences, Wageningen University, The Netherlands.

Legrand, J. (1992), “The Theory of Government Failure”, *British Journal of Political Science,* Vol. 21, pp. 423-42.

Löthgren, M. and M. Tambour (1997), “Bootstrapping the Data Envelopment Analysis Malmquist Productivity Index”, *Applied Economics,* Vol. 31, No. 4, April.

Maniadakis, N., B. Hollingsworth and E. Thanassoulis (1999), “The Impact of the Internal Market on Hospital Efficiency, Productivity and Service Quality”, *Health Care Management Science*, Vol. 2, No. 2, pp. 75-85.

Mueller, D.C. (1989), *Public Choice II*, Cambrige, Cambrige University Press.

Musgrave, R.A. (1981), “Leviathan Cometh: Or Does He?”, in H.F. Ladd and T.N. Tideman (eds.), *Tax and Expenditures Limitations*, Washington (D.C.), Urban Institute Press.

Niskanen, W.A. (1971), *Bureaucracy and Representative Government*, Chicago and New York, Aldine Atherton.

Pedraja-Chaparro, F., J. Salinas-Jiménez and P. Smith (1997), “On the Role of Weight Restrictions in Data Envelopment Analysis”, *Journal of Productivity Analysis,* Vol. 8, No. 2, pp. 215-30.

————— (1999), “On the Quality of the Data Envelopment Analysis”, *Journal of the Operational Research Society.*

Pestieau, P. and H. Tulkens (1990), “Assessing the Performance of Public Sector Activities: Some Recent Evidence from the Productive Efficiency Viewpoint”, Discussion Paper, No. 9060, November, Center for Operations Research & Econometrics, Université Catholique de Louvain.

Ray, S.C. (1991), “Resource Use Efficiency in Public Schools: A Study of Connecticut Data”, *Management Science*, Vol. 37, No. 12, pp. 1620-28.

Salinas-Jiménez, J., F. Pedraja-Chaparro and P. Smith (2003), “Evaluating the Introduction of a Quasi-market in Community Care: Assessment of a Malmquist Index Approach”, *Socio Economic Planning Science* (2003), Vol. 37, No. 1, pp. 1-13.

Sexton, T.R. (1986), “The Methodology of Data Envelopment Analysis”, in R. Silkman (ed.), *Measuring Efficiency: An Assessment of Data Envelopment Analysis*, San Francisco, Jossey Bass.

Smith, P. and M. Goddard (2003), “Los Indicadores de Gestión en el Sector Público: Fortalezas y Debilidades”, Papeles de Economía Española, No. 95, pp. 35-47.

Simar, L. and P.W. Wilson (1998), “Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models”, *Management Science*, Vol. 44, No.1, pp 49-61.

———— (2003), "Estimation and Inference in Two-stage, Semiparametric Models of Production Processes", Discussion Paper, No. 0307, Institut de Statistique, Université Catholique de Louvain.

Stiglitz, J.E. (1989), *The Economic Role of the State*, Basil Blackwell, (v. c. Instituto de Estudios Fiscales, 1993).

Smith, P. (1997), "Model Misspecification in Data Envelopment Analysis", *Annals of Operations Research*, Vol. 73, pp. 233-52.

Tulkens, H. (1993), "On FDH Efficiency Analysis: Some Methodological Issues and Applications to Retail Banking, Courts and Urban Transit", *Journal of Productivity Analysis,* Vol. 4, No. 1-2, pp. 183-210.

Valdmanis, V. (1992), "Sensitivity Analysis for DEA Models – An Empirical Example Using Public Vs. NFP Hospitals", *Journal of Public Economics,* Vol. 48, pp. 185-205.

Wilson, P.W. (1989), *Bureaucracy*, New York, Basic Books.

———— (1995), "Detecting Influential Observations in Data Envelopment Analysis", *Journal of Productivity Analysis*, No. 6, pp. 27-45.

Wilson, P.W. and L. Simar (1995), "Bootstrap Estimation for Non-parametric Efficiency Estimates", Discussion Paper, No. 9517, Institut de Statistique, Université Catholique de Louvain.

Wolf, C. (1979), "A Theory of Non-market Failure", *Journal of Law and Economics,* Vol. 22, No. 1, pp. 107-39.

———— (1987), "Market and Non-Market Failures: Comparison and Assessment", *Journal of Public Policy,*. Vol. 7, No.1, pp. 43-70.

———— (1988), *Markets or Governments: Choosing Between Imperfect Alternatives*, Cambridge (Mass.), M.I.T. Press.