Eficiencia productiva y fronteras de producción. Una aplicación al sector eléctrico

Luis R. Murillo-Zamorano*

Departamento Economía Aplicada y Organización de Empresas, Universidad de Extremadura, Avda. de Elvas s/n , 06071 Badajoz

RESUMEN

Desde que M.J.Farrell publicara en 1957 su artículo "The measurement of productive efficiency", la medición de la eficiencia productiva ha estado intímamente ligada a la construcción de fronteras de producción. Diferentes técnicas han sido utilizadas en la literatura económica a tal efecto. Este trabajo presenta una amplia gama de ellas, tanto parámetricas como no paramétricas. Dichas técnicas son aplicadas a una muestra de 70 empresas eléctricas norteamericanas. El trabajo analiza igualmente la robustez de las citadas técnicas, mediante la comparación de los rankings de ineficiencia derivados de cada una de ellas. Finalmente, del análisis de la estructura productiva y de la estructura de mercado de las empresas objeto de estudio se derivan una serie de factores que pueden favorecer la existencia de una mayor o menor ineficiencia productiva en el sector eléctrico.

Palabras clave: Eficiencia productiva, fronteras paramétricas, DEA, sector eléctrico.

* Tel: 924-289300-ext.9183, fax: 924-272509, e-mail: lmurillo@unex.es;

1. INTRODUCTION

Since such authors as Debreu (1951), Koopmans (1951) or Farrell (1957) introduced the analysis of efficiency in the economic literature, where has been a numerous and wide ranging collection of papers and articles devoted to the measurement of productive efficiency. There has always been a close link between the measurement of efficiency and the use of frontier functions. Different techniques have been utilised to either calculate or estimate these efficient frontiers. The aim of this paper is to provide new insights on their joint use and their application to an industrial organisation framework.

Most of the papers related to the measurement of productive efficiency have based their analysis either on parametric or on non-parametric methods. The choice of estimation method has been an issue of debate, with some researchers preferring the parametric approach (e.g. Berger, 1993) and others the non-parametric approach (e.g., Seiford and Thrall, 1990). The main disadvantage of non-parametric approaches is their deterministic nature. Data Envelopment Analysis (DEA), for instance, does not distinguish between technical inefficiency and statistical noise effects. On the other hand, parametric frontier functions demand the definition of a specific functional form for the technology and for the inefficiency error term. The functional form requirement causes both specification and estimation problems. Obviously, it would be desirable to introduce more flexibility into the parametric frontiers, as well as more thoroughly investigate the non-parametric and stochastic methodologies (e.g. Sengupta, 1987). In our opinion neither approach seems to be strictly preferable. Instead, we think that the joint use of the two groups of techniques can improve the accuracy with which they measure productive efficiency. Following recent literature (e.g., Sengupta, 1995), the main contribution of this paper is to offer some mechanisms and conditions under which that collaboration can be successful.

The data set utilised is partially taken from the one used in Lee (1995). The paper of Lee examines the issue of vertical integration in the US electricity industry in 1990. Three stages --generation, transmission, and distribution-- are analysed in his study. Our study focuses just on the generation stage and therefore no comparative analysis with Lee's study is made.

We organise the paper as follows. Section 2 introduces the techniques used to measure the productive efficiency. Section 3 presents the data set and discusses the results. The fourth section concludes the paper.

2. METHODS

2.1. The Parametric Approach

The parametric approach is naturally subdivided into deterministic and stochastic models. Deterministic models envelope all the observations, identifying the distance between the observed production and the maximum production, defined by the frontier and the available technology, as technical inefficiency. On the other hand, stochastic approaches permit one to distinguish between technical efficiency and statistical noise.

The measurement of productive efficiency by means of parametric techniques requires the specification of a particular frontier function. The Duality theory suggests the use of cost functions to define the production structure. Nerlove (1963) introduced the use of cost functions in the analysis of regulated industries with his application to electric sector. The output produced by firms under a regulated environment, as well as the prices they pay for factors in competitive markets, can be considered to be exogenous. This fact makes the choice of cost functions attractive.

Every cost function implies a set of derived demand equations. Christensen and Greene (1976) argued that the joint use of a cost function and a set of cost share equations as a multivariate regression system provides better estimates of the production structure than those derived from single equation procedures. The dual frontier econometric approach has also evolved from the estimation of single cost functions (e.g., Greene, 1990) to multiple equation systems (e.g., Ferrier and Lovell, 1990; Kumbhakar, 1991). However, some serious estimation and specification problems first noted by Greene (1980), and Nadiri and Schankerman (1981), still remain unsolved. This fact has oriented us to the specification of a single production function.[1]

The technology form adopted was a Cobb-Douglas production function. The frontier production function finally specified can be represented as

$$\log Y_i = a + \sum_{k=1}^{r} b_k \log X_{k,i} + v_i - u_i \tag{1}$$

where i=1,...N indicates the units, Y_i is output, $X_{k,i}$ are inputs. The term $v_i - u_i$ is the composed error term where v_i represents randomness (or statistical noise) and u_i represents technical inefficiency. In the deterministic approach v_i will equal zero.

Several techniques have been developed in the econometric literature in order to estimate deterministic frontier models. In Corrected Ordinary Least Squares (COLS)[2] methodology, the model's parameters, except the intercept term, can be consistently estimated by Ordinary Least Squares (OLS) since that estimation procedure is robust to non-normality[3]. If the estimated intercept term is corrected by shifting it upward until no residual is positive and at least one is zero, we also get a consistent estimator of the intercept term.

Let us assume the following model:

$$y_i = a + \sum_j b_j X_{ij} + e_i \qquad \text{where } e_i \sim \mathrm{N}(0,\sigma^2)$$

Thus,

$$\hat{b}_{j_{COLS}} = \hat{b}_{j_{OLS}}$$

$$\hat{a}_{COLS} = \hat{a}_{OLS} + \max \hat{e}_i$$

$$\hat{m}_{i_{COLS}} = \hat{e}_i - \max \hat{e}_i \tag{2}$$

[1] Panel data techniques can also improve the accuracy of the parametric approach to the measurement of productive efficiency. For a detailed comparative analysis of these techniques, see Kumbhakar (1997).
[2] Gabrielsen (1975).
[3] This was first noted by Richmond (1974).

and individual technical efficiency will be

$$TE_i = e^{-\hat{m}_{iCOLS}}$$

Unlike the deterministic approach, the stochastic frontier models[4] capture the effects of exogenous shocks beyond the control of the analysed units. Errors in the observations and in the measurement of output are also taken into account in this kind of models.

For the Cobb-Douglas case, the stochastic frontier can be represented by eq. (1). The error representing statistical noise is assumed to be identical independent and identically distributed. With respect to the one-sided (inefficiency) error, a number of distributions have been assumed in the literature, being the most frequently used half-normal (SFN), truncated from below at zero (SFT) and exponential (SFE). If the two error terms are assumed independent of each other and of the input variables and some of the previous distributions is used, then the likelihood functions can be defined and maximum likelihood estimates can be determined.

Once the model has been estimated by using maximum likelihood techniques, we obtain a fitted value for the composed error term $v_i - u_i$. For efficiency measurement, we need to separate these two error terms. Jondrow, Lovell, Materov and Schmidt (1982) proposed one way to do it. They developed an explicit formula for the expected value of u_i conditional on the composed error term ($E(u_i \mid v_i - u_i)$) in the half-normal and exponential cases.

Half-normal case:

$$E[u_i|e_i] = \frac{s\,l}{(1+l^2)}\left[\frac{f(e_i\,l\,/\,s\,)}{\Phi(-e_i\,l\,/\,s\,)} - \frac{e_i\,l}{s}\right] \qquad (3)$$

where *f(.)* is the density of the standard normal distribution and *F(.)* the cumulative density function.

---

[4] Aigner, Lovell and Schmidt (1977), Meeusen and van den Broeck (1977), and Battese and Corra (1977).

Exponential case:

$$E\left[u_i|e_i\right] = (e_i - q s_v^2) + \frac{s_v f\left[(e_i - q s_v^2)/s_v\right]}{\Phi\left[(e_i - q s_v^2)/s_v\right]} \tag{4}$$

where $q = 1/s_u$.

Truncated case:

Greene (1993) shows that the conditional technical inefficiencies for the truncated model are obtained by replacing $e_i l/s$ in the expression for the half-normal case, with

$$u_i^* = \frac{e_i l}{s} + \frac{u_i}{s l} \tag{5}$$

Finally, individual (conditioned) technical efficiency scores will be

$$TE_i = e^{-E[u_i|e_i]}$$

2.2. The Non-parametric Approach

Non-parametric analysis (Charnes, Coopers and Rhodes, 1978) does not require the specification of any particular functional form to describe the efficient frontier or envelopment surface. The flexibility of non-parametric techniques allows for several alternative formulations. In this paper we analyse two versions of an output-oriented DEA model according to which returns hypothesis is assumed: namely, constant returns to scale (DEAc) and variable returns to scale (DEAv).

Consider a set of *n* homogenous Decision Making Units (DMU). There are *m* inputs and *s* outputs and each DMU is characterised by an input-output (X,Y) vector. In order to determine the efficiency score of each unit, these will be compared with a peer group consisting of a linear combination of efficient DMUs. For each unit not located on the efficient frontier we define a vector $\overline{m} = (m_1, \ldots, m_n)$ where each m_j represents the weight of each DMU within that peer group. The DEA calculations are designed to maximise the relative efficiency score of each unit, subject to the constraint that the set of weights obtained in this manner for each DMU must also be feasible for all the others included in the sample. That efficiency score can be calculated by means of the following mathematical programming formulation where technical efficiency scores will be determined by the optimum y. Constant (TEc) and variable returns to scale (TEv) formulations are described.

$$
\begin{array}{llll}
TE_C = max_m y^0 & & TE_V = max_m y^0 & \\
s.t. & & s.t & \\
\sum_{j=1}^{n} m_j Y_{ij} \geq y Y_i^0 & \mathrm{i = 1,\ldots,m} & \sum_{j=1}^{n} m_j Y_{ij} \geq y Y_i^0 & \mathrm{i = 1,\ldots,m} \\
\sum_{j=1}^{n} m_j X_{rj} \leq X_r^0 & \mathrm{r = 1,\ldots,s} & \sum_{j=1}^{n} m_j X_{rj} \geq X_r^0 & \mathrm{r = 1,\ldots,s} \\
. & & \sum_{j=1}^{n} m_{\mathrm{j}} = 1 &
\end{array}
\tag{6}
$$

DEA can also be used to calculate scale efficiency. Total technical efficiency is defined[5] in terms of equiproportional increases in outputs that the firm could achieve while consuming the same quantities of its inputs if it were to operate on the constant returns to scale (CRS) production frontier. Pure technical efficiency measures the increase in outputs that the firm could achieve if it were to use the variable returns to scale (VRS) technology. Finally, scale efficiency would be calculated as the ratio of total technical efficiency to pure technical efficiency. If scale efficiency equals one, the firm is operating at CRS, otherwise it would be characterised by VRS[6].

3. DATA AND RESULTS

A wide range of papers related to the treatment of the electric sector with frontier techniques is available in the empirical literature. Schmidt and Lovell (1979,

[5] According to an output-oriented model formulation.

1980) and Greene (1990) introduced the analysis of electricity sector data sets into frontier functions literature. Fare, Grosskopf and Logan (1985) utilise mathematical programming techniques to calculate six different measures of efficiency and compare public versus private performance of electric utilities. Hjalmarsson and Veiderpass (1992) study the local retail distribution of electricity in Sweden in 1985. They apply different versions of the DEA model to 329 firms. Using DEA techniques and OLS analysis, Pollit (1994) examines the cost efficiency in 129 electricity transmission and 145 electricity distribution systems in 1990. Lastly, Ray and Mukerjee (1995) perform a comparative analysis of parametric frontier dual cost functions and non-parametric techniques applied to the data set used previously in Greene (1990).

The data set used in the present empirical application corresponds to a sample of 70 US (investor-owned) electric utility firms in 1990. These firms are approximately evenly spread across the United States. Table 1 provides descriptive statistics for each as used in this study.

<<< TABLE 1 >>>

The capital stock variable is constructed for four different asset classes: steam, nuclear, hydroelectric and other power-generating equipment. In any case, steam technology counts for most of the electricity generated by the companies analysed in this study. The labour variable indicates the number of workers of each firm. There are four main categories of fuel: coal, oil, natural gas, and nuclear. BTU equivalents are used to aggregate different types of fuels over all plants belonging to one firm. The fuel variable is measure in millions of BTUs used in generation of electricity. Finally, total output is indicated in megawatts hours (MWh).

3.1 Efficiency Scores

With respect to the parametric frontiers the estimated parameters of the deterministic and stochastic production functions are given in table 2.

---

[6] Whether those variable returns to scale represent increasing or decreasing returns to scale will depend on the relationships among technical efficiency scores calculated under constant, variable or non-increasing returns to scale.

<<< TABLE 2 >>>

These results come from estimating eq. (1) by means of COLS and MLE, where i=1,...70 indicates the firms, Y_i the output, $X_{1,i} = K_i$ the Capital stock, $X_{2,i} = L_i$ the number of workers, and $X_{3,i} = F_i$ the fuel; b_1, b_2 and b_3 are the elasticities of output with respect to capital, labour and fuel. We infer the presence of constant returns to scale in all the specifications analysed[7]. We estimate a Cobb-Douglas production function. More flexible technologies, such as different versions of translog production functions, presented major problems in the significance of their estimated parameters. Without the factor share equations, estimation of full translog functions can be hampered by an important problem of multicollinearity.[8]

Each of the stochastic specifications yields similar estimates for the partial elasticities of output with respect to capital, labour and fuel. This result seems to confirm the robustness of the technology and distribution hypotheses assumed in the specification of the model.

Table 3 reports the average technical efficiency measures for each of the models explained in the Methods section.[9]

<<< TABLE 3 >>>

As the theory advances, the average efficiency scores of parametric deterministic techniques are lower than the ones estimated through stochastic frontier approaches. Given that COLS is a not stochastic procedure, noise is also reported as inefficiency.

COLS shifts all the residuals down to non-positive values and only one firm of the sample is estimated as efficient[10]. With respect to the DEA approaches, given that

---

[7] Actually, this hypothesis was strongly accepted when we imposed the constraint (b_1) + (b_2) + (b_3) = 1 to the initially unrestricted model. The estimation procedure was made using Limdep 7.0.

[8] According to *Klein's rule of thumb,* multicollinearity is a problem if max $R_j^2 > R^2$ where R_j^2 is the R^2 statistic from the OLS estimation of the auxiliary regression of the j[th] regressor on the other regressor and the intercept term. Several auxiliary regressions were estimated and in all of them this condition was found. Moreover, when we checked the functional form specification of the model, applying a RESET-Test, the Cobb-Douglas technology turned out to be well specified.

[9] The individual efficiency scores generated by each method are given in Appendix 1.

[10] The one with the largest positive OLS residual.

the constraint set is less restrictive under CRS than under VRS, lower efficiency scores are reported for the former case. In our example, DEAc presents an average level of technical efficiency of 73.32% while DEAv efficiency average is 78.71%. For the same reason, fewer units are found to be efficient under CRS than under VRS.

Within the stochastic approaches, no noticeable differences arise. The average efficiency is lower with normal/half-normal models than with the normal/exponential or normal/truncated models, but, in any case, the choice of distribution assumptions does not seem to have a significant effect on the values of the efficiency estimates.

Stochastic frontier models' estimates of σ_v^2 and σ_u^2 provide us with a measure for the relative importance of statistical noise and inefficiency in the estimation of frontier production functions. The variance of the composed error term σ_e^2 is defined as the sum of the variance of the inefficiency error term σ_u^2 and the variance of the statistical noise term σ_v^2. Therefore the participation (%) of each of these components - u and v - in the aggregated error term e can be determined by means of the relationships $\%_u = \sigma_u^2 / (\sigma_u^2 + \sigma_v^2)$ and $\%_v = \sigma_v^2 / (\sigma_u^2 + \sigma_v^2)$. According to the information in table 2, noise represents 59.72% of total variance in the exponential model. In the half-normal and in the truncated cases, these proportions are lower, 25.18% and 17.08% respectively, but still broadly indicative of the importance of noise in the estimation of these models. Therefore, the fact that deterministic models do take noise into account seems to be quite important in our illustrative application. Especially noticeable is the COLS procedure where the average level of technical efficiency is around 60%. These models therefore suffer from both drawbacks: the problems of a rigid specification associated to their parametric nature, and the shortcoming of not distinguishing between inefficiency and noise given their deterministic structure.

3.2. Robustness

Having analysed the efficiency scores, we explore the consistency of the above models in ranking the 70 electric utilities that make up our sample. We are interested in the robustness of the relative position of each electric utility to the use of different methods, rather than in the average levels of technical efficiency found. Table 4 presents pairwise Spearman rank correlation coefficients of the efficiency scores yielded by the six methods used in our analysis[11].

<<< TABLE 4 >>>

These results show that parametric models are extremely consistent in ranking the units. Their pairwise correlation coefficients are not less than 99%. The correlation is also high between parametric techniques and DEAc. On the other hand, correlation coefficients between DEAv and both the econometric approaches and DEAc are not so high. They are around 83% for the group of parametric techniques and 89% for the DEAc model. All parametric approaches were also estimated by imposing the CRS constraint. It seems that the choice of parametric or non-parametric techniques, deterministic or stochastic approaches, or between different distribution assumptions within stochastic techniques is irrelevant if one is interested in ranking electric utilities according to their individual efficiency scores. Only the VRS specification leads to certain differences in those rankings, although such differences are not so large as to stop these rankings still being comparable with the others.

From the calculated technical efficiency scores for each electric using both the CRS and the VRS assumptions, scale efficiency for each unit can be calculated as the ratio TEc/TEv. In our case, the average scale efficiency[12] is 91.40 %, so that, the average inefficiency due to scale reasons of scale is just 8.60 %.

There is also detect an almost perfect correlation between the size of the efficient firms and their returns to scale, in the sense that the bigger firms have decreasing returns to scale and vice versa. It seems that economies of scale are exhausted at the greatest levels of production while they are still available at lower levels. This result agrees with the low value found for the average scale inefficiency and is supporting evidence that the units in our sample are operating at the correct scale. Some studies as Cummins and Zi (1997), for example, have found a direct relationship between the size of units and their inefficiency levels. In our case, no such relationship seems to appear.

[11] Spearman's correlation coefficients were calculated using the SPSS 8.0 package.
[12] Scale efficiency levels for each of the observations in the sample are given in Appendix 2.

So far, we have analysed different methods and their robustness in the measurement of productive efficiency. The next step in this empirical application will provide some possible explanations for the efficiency scores described above.

3.3. Inefficiency Sources

One common practice in the literature is to regress the efficiency scores against a vector of explanatory variables. Disaggregated data for different types of capital and output are used as proxies for the productive structure and market demand structure faced by each electric utility. Capital stock levels attached to steam, nuclear and hydroelectric assets are used to evaluate the influence of each of those technologies on higher or lower efficiency scores. Similarly, the allocation of total megawatt-hours to three different demand categories -- commercial, industrial and residential -- is also considered on the basis of explaining individual efficiency scores.

The high degree of correlation between those proxies for productive and market structure and the original variables specified in our model is a handicap for two-stage models. However, the choice of a one stage model, as Lovell (1993) points out does not solve this problem of correlation between the variables used in the initial specification of the model and those used in the subsequent analysis of the efficiency sources: it just replaces a problem of omitted (two stages model) with one of multicollinearity.[13]

For the series of inefficiency scores to take into account as the dependent variable, we have used that generated by the DEAc model[14]. The DEA-based efficiency scores are truncated from below at one. OLS regression would produce biased and inconsistent parameter estimates, so we use a truncated regression model (Tobit model). The estimated parameters are given in table 5.

<<< TABLE 5 >>>

[13] Some functional forms with dissaggregated levels of capital and output used as regressors were also estimated. However, such a large list of variables, especially in the translog version, and the high degree of correlation among them requires a very high order in the convergence criteria of the maximum likelihood algorithms of stochastic frontier models. This precluded the estimation of these stochastic models.

[14] The results with the COLS, SFN, SFE and SFT efficiency series were almost identical.

Given the statistical significance of the three parameters used as proxies, it seems that the productive structure affects the efficiency scores attained by the different electric utilities. The market demand structure, on the other hand, seems not to have any influence.

The variables used to measure the effects of market demand structure on the inefficiency of each unit are characterised by a high degree of homogeneity across observations (see table 1). Therefore it is not surprising to find that they are not significant explanations for the inefficiency of units.

Within productive structure factors, steam and nuclear technologies are found to be directly related to inefficient behaviour of the units in the sample, while the use of hydroelectric technology seems to have positive effects on their efficincy. Nuclear and even more so steam technologies seem to be exhausting their particular economies of scale.

The main problem of "two-stage" models, such as that used in this paper, is to know which regressors must be included in the estimation of efficiency levels and which in their explanation. In the light of our results, besides their not being highly correlated with the variables utilised in the frontier estimation procedure, a necessary although not sufficient condition for regressors to be considered as proxies for inefficiency sources is that they must be able to introduce heterogeneity in the analysis. Thus, a necessary extension to the empirical analysis that we have so far presented would be the introduction of additional information through variables properly representative of the industrial organisation, such as market structure, regulatory environment, ownership or internal organisation of the firm.

4. CONCLUSIONS

The main problem of the "two-stage" models as the ones used in this paper is to know what regressors must be included in the estimation of efficiency levels and which others in their explanation. Looking at our results, besides the fact of not being highly correlated with the variables utilised in the frontier estimation procedure, a necessary although not sufficient condition for regressors be considered as proxies of inefficiency sources is that they must be able to introduce heterogeneity in the analysis. So, a necessary extension to the empirical analysis so far presented would be the introduction of additional information through variables properly representative of factors as market structure, regulatory environment, ownership or internal organisation of the firm.

In accordance with the Spearman rank correlation coefficients, only the DEA_V introduces some differences in the efficiency rankings. However, they can still be considered as comparable ones. Moreover, no relationship between inefficiency and the size of the electricity supply utilities is founded in our study.

Finally, the consistency of the models, either parametric or non-parametric, in ranking inefficient untis, confirms our opinion that no technique is strictly better than other. Each one has its own advantages and disadvantages. A careful consideration of them, of the data set utilised, and of the intrinsic characteristics of the industry under analysis will help us in the correct implementation of these techniques.

REFERENCES

Aigner, D.J., C.A.K. Lovell and P.J. Schmidt, 1977, Formulation and estimation of stochastic frontier production function models, Journal of Econometrics 6, 21-37.

Banker, R.D., and R. Morey, 1986a, Efficiency analysis for exogenously fixed inputs and outputs, Operations Research 34 (4), 513-521.

Banker, R.D., and R. Morey, 1986b, The use of categorical variables in data envelopment analysis, Management Science 32 (12), 1613-1627.

Battese G. and G. Corra 1977. Estimation of a production frontier model with application to the pastoral zone of Eastern Australia, Australian Journal of Agricultural Economics 21 (3), 167-179.

Berger, A.N. 1993, Distribution-free estimates of efficiency in the U.S. banking industry and tests of the standard distributional assumptions, Journal of Productivity Analysis 4, 261-292.

Cummins, J.D., and H. Zii, 1997, Measuring cost efficiency in the U.S. life insurance industry: econometric and mathematical programming approaches, Working paper 97-03, The Wharton School. University of Pennnylvania.

Charnes, A., W.W. Cooper and E. Rhodes 1978, Measuring the efficiency of decision-making units, European Journal of Operational Research 2, 429-444.

Charnes, A., W.W. Cooper, Q.L. Wei, and Z.M. Huang, 1989, Cone ratio data envelopment analysis and multi-objective programming, International Journal of Systems Science 20 (7), 1099-1118.

Christensen, L.R. and W.H. Greene, 1976, Economies of scale in U.S. electric power generation, Journal of Political Economy 84 (4), 655-676.

Debreu, G., 1951, The coefficient of resource utilization, Econometrica 19, 273-292.

Fare, R., S. Grosskopf and J.Logan, 1985, The relative performance of publicily-owned and privately-owned electric utilities, Journal of Public Economics 26, 89-106.
Farrell, M.J., 1957, The measurement of productive efficiency, Journal of the Royal Statistical Society (A, general) 120, 253-281.
Ferrier, G.D. and C.A.K. Lovell, 1990, Measuring cost efficiency in banking: econometric and linear programming evidence, Journal of Econometrics 46, 229-245.
Gabrielsen, A., 1975, On estimating efficient production functions, Working Paper no A-35, Chr. Michelsen Institute, Department of Humanities and Social Sciences, Bergen, Norway.
Greene, W., 1980, On the estimation of a flexible frontier production model, Journal of Econometrics 13 (1), 101-115.
Greene, W., 1990, A gamma-distributed stochastic frontier model, Journal of Econometrics 46, 141-163.
Greene, W.M., 1993, The econometric approach to efficiency analysis, in: Harold O. Fried, C.A.K. Lovell and S.S. Schmidt, eds., The Measurement of Productive Efficiency: Techniques and Applications (Oxford University Press) 68-119.
Hjalmarsson, L. and A. Veiderpass, 1992, Efficiency and ownership in Swedish electricity retail distribution, The Journal of Productivity Analysis 3, 7-23.
Jondrow, J, C.A.K. Lovell, I.S. Materov and P. Schmidt, 1982, On the estimation of technical inefficiency in the stochastic frontier production function model, Journal of Econometrics 23, 269-274.
Koopmans, T.C., 1951, An analysis of production as efficient combination of activities, in: T.C. Koopmans, ed., Activity Analysis of Production and Allocation (Cowles Commission for Research in Economics, NewYork).
Kumbhakar, S.C., 1991, The Measurement and Decomposition of Cost-Inefficiency: The Translog Cost System, Oxford Economic Papers 43, 667-683.
Kumbhakar, S.C., A. Heshmati and L. Hjalmarsson , 1997, Temporal patterns of technical efficiency: Results from competing models, International Journal of Industrial Organization 15(5), 597-616.
Lee, B-J, 1995, Separability test for the electricity supply industry, Journal of Applied Econometrics 10, 49-60.
Lovell, C.A.K., 1993, Production frontiers and productive efficiency, in: O. Harold Fried, C.A.K. Lovelland S.S. Schmidt, eds., The Measurement of Productive Efficiency: Techniques and Application (Oxford University Press) 3-67.
Maddala, G.S., 1983, Limited dependent and qualitative variables in econometrics (Cambridge University Press, New York).
Meeusen, W. and J. van den Broeck, 1977, Efficiency Estimation from Cobb-Douglas production functions with composed error, International Economic Review 18, 435-444.
Nadiri, M.I. and M.A. Schankerman, 1981, The structure of production, technological change, and the rate of growth of total factor productivity in the US bell system, (Academic Press, New York).
Nerlove, M., 1963, Returns to scale in electricity supply (Stanford University Press, Stanford).
Pollit, M.G., 1994, Productive efficiency in electricity transmission and distribution systems, Oxford Applied Economics Discussion Paper Series 161.
Ray, S.C., and K. Mukherjee, 1995, Comparing parametric and nonparametric measures of efficiency: a reexamination of the Christensen-Greene data, Journal of Quantitative Economics 11 (1), 155-168.
Richmond, J., 1974, Estimating the efficiency of production, International Economic Review 15, 515-521.

Schmidt, P. and C.A.K. Lovell, 1979, Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers, Journal of Econometrics 9, 343-366.
Schmidt, P. and C.A.K. Lovell, 1980, Estimating stochastic production and cost frontiers when technical and allocative inefficiency are correlated, Journal of Econometrics 13, 83-100.
Seiford, L.M., and R.M. Thrall, 1990, Recent development in DEA: the mathematical programming approach to frontier analysis, Journal of Econometrics 46, 7-38.
Sengupta, J.T., 1987, Data envelopment analysis for efficiency measurement in the stochastic case, Computers and Operations Research 14, 117-169.
Sengupta, J.T., 1995, Estimating efficiency by cost frontiers: a comparison of parametric and nonparametric methods, Applied Economics Letters 2, 86-90.
Timmer, C.P., 1971, Using a probabilistic frontier function to measure technical efficiency, Journal of Political Economy 79, 579-597.

APPENDIX 1

obs.	TECOLS	TESFN	TESFT	TESFE	TEDEAC	TEDEAV
1	0.3914	0.6233	0.6918	0.6729	0.5251	0.6181
2	0.5068	0.7681	0.8632	0.8504	0.7679	1.0000
3	0.5376	0.7887	0.8601	0.8619	0.6387	1.0000
4	0.6405	0.8666	0.9111	0.9138	0.8249	1.0000
5	0.2895	0.4973	0.5434	0.4967	0.3330	0.3492
6	0.4994	0.7556	0.8320	0.8340	0.5941	0.5979
7	0.4686	0.7180	0.7964	0.7938	0.5551	0.5772
8	0.7368	0.9116	0.9366	0.9396	0.9583	1.0000
9	1.0000	0.9486	0.9632	0.9593	1.0000	1.0000
10	0.6115	0.8467	0.8996	0.9008	0.7947	0.8614
11	0.4800	0.7303	0.8101	0.8061	0.5604	0.5670
12	0.5038	0.7574	0.8357	0.8333	0.5863	0.6077
13	0.6495	0.8743	0.9178	0.9182	0.7724	0.7749
14	0.5539	0.8008	0.8772	0.8689	0.6522	0.7012
15	0.8360	0.9354	0.9514	0.9526	1.0000	1.0000
16	0.4889	0.7567	0.8295	0.8361	0.5919	0.5926
17	0.7961	0.9237	0.9485	0.9456	0.9884	1.0000
18	0.5166	0.7689	0.8427	0.8415	0.6490	0.6534
19	0.5206	0.7981	0.8851	0.8748	0.8762	0.9199
20	0.5656	0.8024	0.8758	0.8666	0.7486	0.7521
21	0.5564	0.8136	0.8766	0.8792	0.6607	0.6883
22	0.6542	0.8819	0.9224	0.9232	0.8228	0.8318
23	0.4935	0.7488	0.8272	0.8238	0.5716	0.6197
24	0.7436	0.9096	0.9384	0.9371	0.8540	0.8976
25	0.6528	0.8702	0.9179	0.9142	0.7290	0.8110
26	0.5581	0.8076	0.8769	0.8731	0.6353	0.7039
27	0.4768	0.7396	0.8117	0.8164	0.5681	0.6009
28	0.8127	0.9175	0.9471	0.9398	1.0000	1.0000
29	0.6471	0.8828	0.9191	0.9237	0.7810	0.7858
30	0.8372	0.9337	0.9509	0.9509	1.0000	1.0000
31	0.4571	0.7225	0.7941	0.7998	0.5513	0.5758
32	0.6113	0.8613	0.9057	0.9108	0.7376	0.7553
33	0.5660	0.8284	0.8834	0.8897	0.6801	0.7188
34	0.5465	0.8067	0.8711	0.8731	0.6451	0.7010
35	0.6227	0.8691	0.9118	0.9154	0.7593	0.7929
36	0.5982	0.8547	0.8986	0.9065	0.7338	0.8248
37	0.5850	0.8383	0.8907	0.8947	0.6988	0.7683
38	0.4366	0.6878	0.7685	0.7548	0.5068	0.5870
39	0.5745	0.8254	0.8819	0.8846	0.8098	0.8214
40	0.5278	0.7832	0.8505	0.8517	0.7462	0.7557

41	0.6239	0.8684	0.9105	0.9144	0.7509	0.7903
42	0.6313	0.8735	0.9133	0.9174	0.7612	0.8042
43	0.5524	0.8176	0.8756	0.8812	0.6609	0.7208
44	0.5431	0.8055	0.8640	0.8708	0.7439	0.8028
45	0.4825	0.7417	0.8146	0.8140	0.5748	0.6434
46	0.8370	0.9400	0.9537	0.9552	1.0000	1.0000
47	0.5145	0.7773	0.8515	0.8486	0.6142	0.7024
48	0.5955	0.8562	0.8972	0.9071	0.7709	0.8889
49	0.7064	0.9023	0.9318	0.9327	0.8320	0.9171
50	0.6028	0.8532	0.9034	0.9043	0.7250	0.8080
51	0.4921	0.7602	0.8284	0.8337	0.5837	0.6589
52	0.6770	0.8863	0.9222	0.9223	1.0000	1.0000
53	0.5388	0.7914	0.8669	0.8575	0.5962	0.7869
54	0.7379	0.9153	0.9395	0.9404	0.8829	0.9667
55	0.6376	0.8730	0.9154	0.9157	0.7492	0.8534
56	0.5766	0.8342	0.8901	0.8912	0.6775	0.7706
57	0.5424	0.8153	0.8727	0.8795	0.6558	0.7332
58	0.6525	0.8806	0.9212	0.9200	0.7819	0.9041
59	0.4991	0.7715	0.8323	0.8430	0.6254	0.7386
60	0.6445	0.8757	0.9159	0.9164	0.7442	0.8550
61	0.6273	0.8618	0.9123	0.9073	0.7073	0.9420
62	0.3204	0.5430	0.5935	0.5500	0.4043	0.5849
63	0.6144	0.8575	0.9081	0.9050	0.7122	0.8696
64	0.7657	0.9189	0.9425	0.9412	0.8874	1.0000
65	0.5856	0.8443	0.8944	0.8969	0.6855	0.8004
66	0.6748	0.8940	0.9261	0.9274	0.7929	0.9096
67	0.7546	0.9211	0.9422	0.9431	0.8948	1.0000
68	0.7107	0.9076	0.9331	0.9348	0.8458	1.0000
69	0.6874	0.8933	0.9261	0.9253	0.9743	1.0000
70	0.6838	0.8959	0.9288	0.9278	0.7841	1.0000

Notes:
obs.: observations ordered by ouput produced.
COLS: Corrected Ordinary Least Squares.
TEcols: Technical efficiency scores with COLS.
SFN: Stochastic Frontier (Half Normal).
TEsfn: Technical efficiency scores with SFN.
SFT: Stochastic Frontier (Truncated).
TEsft: Technical efficiency scores with SFT.
SFE: Stochastic Frontier (Exponential).
TEsfe: Technical efficiency scores with SFE.
DEAc: Data Envelopment Analysis (Constant Returns to Scale).
TEdeac: Technical efficiency scores with DEAc.
DEAv: Data Envelopment Analysis (Variable Returns to Scale).
TEdeav: Technical efficiency scores with DEAv.

APPENDIX 2

obs.(*)	Total Output	TEDEAc	TEDEAv	Scale	Ef.Returns
1	1.678.385.600	0,5251	0,6181	0,8495	
2	1.731.357.000	0,7679	1,0000	0,7679	IRS
3	1.823.194.900	0,6387	1,0000	0,6387	IRS
4	2.382.022.100	0,8249	1,0000	0,8249	IRS
5	2.407.029.100	0,3330	0,3492	0,9536	
6	2.413.488.700	0,5941	0,5979	0,9936	
7	2.652.219.700	0,5551	0,5772	0,9617	
8	2.683.449.600	0,9583	1,0000	0,9583	IRS
9	3.240.713.200	1,0000	1,0000	1,0000	CRS
10	3.312.763.500	0,7947	0,8614	0,9226	
11	3.438.851.100	0,5604	0,5670	0,9884	

IRS: Increasing Returns to Scale
CRS: Constant Returns to Scale
DRS: Decreasing Returns to Scale

TEDEAc

Mean	0,7333
Standard Error	0,0177
Median	0,7408

12	3.842.686.400	0,5863	0,6077	0,9648	
13	4.043.033.000	0,7724	0,7749	0,9968	
14	4.399.966.000	0,6522	0,7012	0,9301	
15	4.473.222.700	1,0000	1,0000	1,0000	CRS
16	4.507.697.000	0,5919	0,5926	0,9988	
17	4.620.701.000	0,9884	1,0000	0,9884	CRS
18	5.465.902.900	0,6490	0,6534	0,9933	
19	5.762.995.000	0,8762	0,9199	0,9525	
20	5.799.669.000	0,7486	0,7521	0,9953	
21	5.807.470.000	0,6607	0,6883	0,9599	
22	5.871.063.000	0,8228	0,8318	0,9892	
23	6.124.283.000	0,5716	0,6197	0,9224	
24	6.205.470.000	0,8540	0,8976	0,9514	
25	6.381.468.000	0,7290	0,8110	0,8989	
26	6.520.492.000	0,6353	0,7039	0,9025	
27	6.964.168.000	0,5681	0,6009	0,9454	
28	7.149.296.000	1,0000	1,0000	1,0000	DRS
29	7.541.245.000	0,7810	0,7858	0,9939	
30	7.721.213.100	1,0000	1,0000	1,0000	CRS
31	8.030.476.000	0,5513	0,5758	0,9575	
32	8.325.322.000	0,7376	0,7553	0,9766	
33	8.602.467.000	0,6801	0,7188	0,9462	
34	9.943.526.000	0,6451	0,7010	0,9203	
35	10.286.717.000	0,7593	0,7929	0,9576	
36	10.397.817.000	0,7338	0,8248	0,8897	
37	10.792.818.000	0,6988	0,7683	0,9095	
38	10.900.298.000	0,5068	0,5870	0,8634	
39	11.507.614.000	0,8098	0,8214	0,9859	
40	11.714.757.000	0,7462	0,7557	0,9874	
41	11.783.880.000	0,7509	0,7903	0,9501	
42	12.399.063.000	0,7612	0,8042	0,9465	
43	13.436.381.000	0,6609	0,7208	0,9169	
44	14.419.022.000	0,7439	0,8028	0,9266	
45	14.779.980.000	0,5748	0,6434	0,8934	
46	15.539.938.000	1,0000	1,0000	1,0000	CRS
47	15.655.908.000	0,6142	0,7024	0,8744	
48	15.722.765.000	0,7709	0,8889	0,8673	
49	16.070.475.200	0,8320	0,9171	0,9072	
50	16.941.931.000	0,7250	0,8080	0,8973	
51	17.687.140.000	0,5837	0,6589	0,8859	
52	19.678.360.000	1,0000	1,0000	1,0000	CRS
53	19.757.191.000	0,5962	0,7869	0,7577	
54	21.436.125.000	0,8829	0,9667	0,9133	
55	21.529.356.000	0,7492	0,8534	0,8779	
56	21.816.200.000	0,6775	0,7706	0,8792	
57	23.340.654.000	0,6558	0,7332	0,8944	
58	25.916.889.000	0,7819	0,9041	0,8648	
59	26.002.927.000	0,6254	0,7386	0,8467	
60	31.603.013.000	0,7442	0,8550	0,8704	
61	32.035.118.000	0,7073	0,9420	0,7508	
62	32.591.836.000	0,4043	0,5849	0,6912	
63	36.192.125.000	0,7122	0,8696	0,8190	
64	36.309.960.000	0,8874	1,0000	0,8874	DRS
65	43.535.926.000	0,6855	0,8004	0,8564	
66	46.868.634.000	0,7929	0,9096	0,8717	
67	51.776.850.000	0,8948	1,0000	0,8948	DRS
68	63.558.870.000	0,8458	1,0000	0,8458	DRS
69	64.410.130.000	0,9743	1,0000	0,9743	DRS
70	70.517.340.000	0,7841	1,0000	0,7841	

Mode	1,0000
Standard Devt	0,1477
Variance	0,0218
Kurtosis	-0,0388
Skewness	0,0202
Range	0,6670
Minimum	0,3330
Maximum	1,0000
Sum	51,327
Count	70,000

TEDEAv

Mean	0,8115
Standard Error	0,0162
Median	0,8028
Mode	1,0000
Standard Devt.	0,1356
Variance	0,0184
Kurtosis	2,4517
Skewness	-1,704
Range	1,0000
Minimum	0,5670
Maximum	1,0000
Sum	51,122
Count	63,000

Scale Efficiency

Mean	0,9140
Standard Error	0,0092
Median	0,9225
Mode	1,0000
Standard Devt.	0,0769
Variance	0,0059
Kurtosis	1,9740
Skewness	-1,2586
Range	0,3613
Minimum	0,6387
Maximum	1,0000
Sum	63,9825
Count	70,0000

Efficient units with CRS: 6
Efficient units with IRS 4
Efficient units with DRS: 6

* Observations ordered
by output produced

TABLES

Table 1. Main descriptive statistics of variables used in the study.

Variable	Mean	Max.	Min.	Standard Deviation
Total Output	15.582	70.517	1.678	1.568E+10
Total Capital	94.914	409.673	9.367	91.747
Total Labour	4.993	24.607	440	5.198
Total Fuel	1.324E+10	4.750E+11	7.001E+09	1.111E+11
% Ksteam (1)	0.7674	0.9999	0.084	0.2192
% Knuclear (1)	0.1120	0.6754	0	0.1762
% Khydroelectric(1)	0.0422	0.3256	0	0.0757
% KOther GE(1)	0.0783	0.9150	0	0.1280
% Ocommercial (2)	0.2664	0.6421	0.037	0.0987
% Oindustrial(2)	0.3485	0.5533	0.1052	0.0774
% Oresidential(2)	0.3850	0.8113	0.063	0.1272

(1) Represents the percentage of capital stock levels attached to steam, nuclear, hydroelectric and other power- generating equipment assets.
(2) Allocation of total MWh to commercial, industrial and residential demand categories.

Table 2.Estimated parameters of deterministic and stochastic production frontiers.
(t-test statistics appear in parentheses)

	COLS	SFN	SFT	SFE
Intercept (a)	10.819(*) (10.014)	11.786 (15.870)	11.145 (13.886)	10.951 (14.453)
Capital (b_1)	0.1392 (2.414)	0.1340 (1.893)	0.1066 (1.330)	0.1391 (2.270)
Labour (b_2)	0.6441 (10.539)	0.6745 (10.865)	0.6713 (10.084)	0.6441 (11.485)
Fuel (b_3)	0.2174 (3.474)	0.1794 (3.954)	0.2170 (4.705)	0.2174 (4.847)
R^2	0.9506			
F	423.529			
Log-Lik.	10.1631	11.3880	11.1224	11.8625
S_u / S_v		1.7239 (1.897)	2.2007 (1.405)	
S_u^2		0.0621	0.0995	0.0176
S_v^2		0.0209	0.0205	0.0261
$\sqrt{S_v^2 + S_u^2}$		0.2881	0.3465	
u / S_u			0.6346	
Theta				7.5318 (2.402)
S_v				0.1617 (5.080)
(b_1) + (b_2) + (b_3)**	1.0007 {0.9756}	0.9879 [0.3570]	0.9949 [0.1421]	1.0006 [0.0230]

(*****) If the estimated intercept term is corrected by shifting it upward until no residual is positive and at least one is zero, we will get a consistent estimator of the intercept term. In our case this consistent intercept is 11.349.
(******) CRS hypothesis test:.{ _ }:Probability associated with an F-Test (1.66). [_]: Significance level in a Wald Test-χ^2 (1).

Table 3. Technical efficiency averages.

Method	Average Efficiency	Max.	Min.	Standard Deviation	Number of efficient units
COLS	60.09	1	28.95	0.123	1
SFN	82.61	94.86	49.72	0.086	0
SFT	87.77	96.31	54.33	0.073	0
SFE	87.64	95.93	49.66	0.080	0
DEAc	73.32	100	33.3	14.77	6
DEAv	78.71	100	6.9	19.39	16

(*) The average efficiency measures of COLS, SFN, SFT, and SFE were estimated under the null hypothesis of Constant Returns to Scale.

Table 4. Spearman correlation coefficients among alternative efficiency measures(*).

	COLS	SFN	SFT	SFE	DEAc	DEAv
COLS	1.000					
SFN	0.994	1.000				
SFT	0.995	0.994	1.000			
SFE	0.991	0.998	0.994	1.000		
DEAc	0.909	0.907	0.918	0.915	1.000	
DEAv	0.833	0.829	0.843	0.835	0.890	1.000

(*) All the correlation coefficients among different methods are significant at the .01 level (2-tailed).

Table 5. Tobit model estimated parameters

Variable (%)	Parameter Estimate	t-student	Mean	Max.	min.	Standard Deviation
Ksteam	0.2975	2.346**	0.7674	0.9999	0.084	0.2192
Knuclear	0.2848	1.856 *	0.1120	0.6754	0	0.1762
KHydro.	-0.4295	-1.820 *	0.0422	0.3256	0	0.0757
Ocommercial	0.1049	0.530	0.2664	0.6421	0.037	0.0987
Oindustrial	0.2526	1.596	0.3485	0.5533	0.1052	0.0774
Oresidential	-0.2848	-1.484	0.3850	0.8113	0.063	0.1272

** Significant coefficients at the 5% level (2-tailed).
* Significant coefficients at the 10% level (2-tailed).