Economic Benchmarking and its Applications

Virendra Ajodhia
Konstantin Petrov
Gian Carlo Scarsi*

1. Introduction

Regulators worldwide are increasingly looking at economic benchmarking techniques to characterise and predict the efficiency levels of the utilities they are called to regulate by law. Regulatory agencies who have to tackle the difficult issue of mimicking market interaction in such industry subsets when natural monopoly is present as transmission and distribution of electricity or gas, generally prefer to adopt rigorous techniques to establish relative efficiency advancements by natural monopolists.

In order to evaluate the efficiency of indirectly competing monopolists, regulators can impose discipline on them by comparing their operations. When done on the basis of economic principles as opposed to process-based engineering checks, this is known as 'economic benchmarking'. Economic benchmarking techniques abound. We will concentrate in this article on those two techniques that, by academic pedigree and practical application, look like the most appealing ones to regulators worldwide, and as such have been applied in a number of jurisdictions.

The paper is strongly relevant for the upcoming energy regulation in Germany. In September this year the Ministry of Economy (BMWA) in Germany presented its findings about the status of the electricity market opening in Germany. Although the report concluded that the Association Agreement was relatively successful and lead to a reasonable quick and effective opening of the German electricity market a number of critical areas have been indicated. Development of an effective network price control mechanism belongs to these areas. The Ministry leaves open, if such price control mechanism shall be similar to the incentive systems developed and implemented in many other countries. However, the new price control mechanism shall ensure that only efficient costs for provision of network service are covered. The application of efficiency assessment techniques and integration of efficiency scores into the price control design is a fundamental question that will need to be addressed by the future German regulator.

The article is structured as follows. Section 2 illustrates the mainly used non-statistical benchmarking technique, Data Envelopment Analysis (DEA). Section 3 describes the statistical counterpart of DEA, which builds up from simple econometric regression analysis with a series of non-trivial complications: this is known as Stochastic Frontier Analysis (Estimation) or SFA/SFE. Section 4 compares the two techniques. Section 5 provides an example of economic benchmarking from the real world, with a focus upon Europe. In Section 6, we present a DEA model used to benchmark cost and quality levels of electricity distribution units in four countries, and discuss the results. Section 7 concludes.

---

* Konstantin Petrov and Gian Carlo Scarsi are with KEMA Consulting, part of KEMA International, and are based in Bonn, Federal Republic of Germany and London, United Kingdom. They are the corresponding authors. Virendra Ajodhia is Associate to KEMA Consulting and a Researcher and Doctoral Student at Delft University of Technology (TU Delft), the Netherlands. Mailing address: KEMA Consulting, Kurt-Schumacher-Str. 8, D-53113 Bonn, Federal Republic of Germany, Tel. +49 228 44 690 00, Fax +49 228 44 690 99; mailto: Virendra.Ajodhia@Kema.com, Konstantin.Petrov@Kema.com, GianCarlo.Scarsi@Kema.com.

2. Microeconomic Efficiency Measurement and DEA: a Theoretical Summary

Economic efficiency analysis is a concept much used in the industrial organisation literature and has its origins in the microeconomic theory of production and cost. The applied use of comparative efficiency analysis is just a by-product of the typical microeconomic problem of the measurement of efficient cost and production levels, and of the separation between different types of inefficiency in production.

This paper outlines a number of commonly used efficiency measures and discusses how they may be calculated relative to an efficient technology, which is generally represented by some frontier function. Frontiers have been estimated by applied economists and econometricians using many different techniques over the past forty years. If we ignore the simple averaged estimation that Ordinary Least Squares (OLS) and its variants provide, the two principal methods[1] which provide some degree of 'best practice' – and not simple central-tendency - outcomes are the following:

(a) Stochastic Frontier Estimation (SFE), also known in the regulatory practice as Stochastic Frontier Analysis (SFA); and

(b) Data Envelopment Analysis (DEA), which is the focus of much regulatory practice in the field of electric utilities.

The two methodologies involve econometric and mathematical programming methods, respectively. The discussion in this Section provides a brief introduction to modern efficiency measurement. A more detailed treatment is provided by Fare, Grosskopf, and Lovell (1985, 1994), and Fried, Lovell, and Schmidt (1993). An interesting overview of DEA is in Seiford and Thrall (1990), whereas the two basic DEA models being developed in the late Seventies and early Eighties - to which most applied papers still refer - are those by Charnes, Cooper, and Rhodes (1978) for constant returns to scale (CRS) DEA, and by Banker, Charnes, and Cooper (1982, 1984) for variable returns to scale (VRS) DEA. A new perspective on DEA in principal-agent theory terms has been provided by Bogetoft (1994) and in more recent applied papers written by this author, sometimes in co-operation with Agrell, on applied regulatory topics in Scandinavia.

Modern efficiency measurement begins with Farrell (1957), who drew upon the work of Debreu (1951) and Koopmans (1951), to define a simple measure of firm efficiency that could account for multiple inputs, and easily generalise to multiple outputs. He claimed that the efficiency of a firm consists of two components:

(*a*) technical efficiency, which reflects the ability of a firm to obtain maximal output from a given set of inputs, and
(*b*) input-allocative efficiency, which reflects the ability of a firm to use the inputs in optimal proportions, given their respective prices.

These two measures are then combined to provide a measure of total economic efficiency[2]. The following discussion begins with Farrell's original ideas, which were illustrated by the author in input-input space and hence had an input-saving focus (i.e., they assumed that firms should minimise input usage for a given set of outputs). These are usually termed input-orientated measures, as opposed to output-maximising (or output-orientated) ones, which assume that firms maximise output(s) for a given set of inputs.

---

[1] See Pollitt (1995).
[2] Some of Farrell's terminology differed from the current one. He used the term *price efficiency* instead of 'input-allocative' efficiency, and the term *overall efficiency* instead of 'economic' efficiency.

Farrell (1957) illustrated his ideas by using a simple example involving firms which utilise two inputs (to make up the input vector, **X**) to produce a single output (**Y**), under the assumption of constant returns to scale (CRS)[3]. Knowledge of the unit isoquant of the fully efficient firm[4], represented by SS′ in Figure 1, permits the measurement of technical efficiency. If a given firm uses quantities of inputs, defined by point P, to produce a unit of output, technical inefficiency for that firm is represented by the distance QP, i.e. the amount by which all inputs could be proportionally reduced without a reduction in output. This is usually expressed by the ratio QP/OP, which represents the percentage by which all inputs could be reduced. The technical efficiency level (TE) of a firm is most commonly measured by the ratio

$$TE_i = OQ/OP, \quad (1)$$

which is equal to one minus QP/OP. It will take a value between zero and one, and hence provides an indicator of the degree of technical (in)efficiency of the firm. A value of one indicates that the firm is fully technically efficient. For example, point Q is technically efficient because it lies on the efficient unit isoquant.

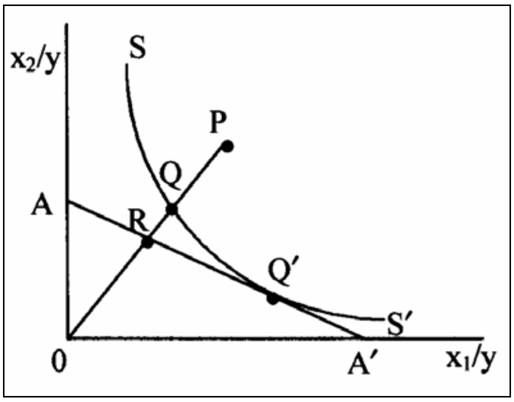

Figure 1: Technical and input-allocative inefficiency. Source: Battese and Coelli (1995).

If the input price ratio, represented by the line AA′ in Figure 1, is also known, allocative efficiency may also be calculated. The allocative efficiency level (AE) of the firm[5] operating at point P is defined by the ratio

$$AE_i = OR/OQ, \quad (2)$$

since the distance RQ represents the reduction in production costs that would occur if production were to take place at the allocatively (and technically) efficient point Q′, instead

[3] The CRS assumption simplifies the analysis by allowing the use of unit isoquants. Furthermore, Farrell also discussed the extension of his method so as to accommodate more than two inputs, multiple outputs, and non-constant returns to scale.
[4] The production function of the fully efficient firm is not known in practice, and thus must be estimated from observations on a sample of firms in the industry concerned. In this paper, DEA is meant to estimate the efficient production, or 'best-practice', frontier.
[5] More generally, data points are for 'decision-making units' (DMUs), which can also be branches of a single company.

of the technically efficient, but allocatively inefficient, point Q[6]. To sum up, the total economic efficiency level (EE) for the *i*-th decision-making unit will be defined as

$$EE_i = OR/OP, \qquad (3)$$

where the distance RP can also be interpreted in terms of a cost reduction. Notice that the product of technical and allocative efficiency provides the overall economic efficiency level (multiplicatively separable), which is also known as 'Farrell efficiency':

$$TE_i \cdot AE_i = \frac{OQ}{OP} \cdot \frac{OR}{OQ} = \frac{OR}{OP} \equiv EE_i. \qquad (4)$$

Notice that all measures are bounded by zero and one, and that overall (Farrell) efficiency has the nice property of multiplicative separability into input-allocative and technical efficiencies. Separability may also be exploited in order to decompose technical efficiency into scale, congestion, and 'purely technical' efficiency, as in Fare et al. (1985). These efficiency measures assume that the production function of the fully efficient, or 'best-practice', firm is known. However, this is rarely the case, so that the best-practice unit isoquant must be estimated from sample data. Farrell suggested the use of either:

(*a*) a non-parametric piecewise-linear convex isoquant constructed such that no observed point should lie to the left or below it (refer to Figure 2), or
(*b*) a parametric function, such as the Cobb-Douglas, CES, generalised Leontief (Diewert), or translog[7] forms being fitted to the data, again such that no observed point should lie to the left or below it.

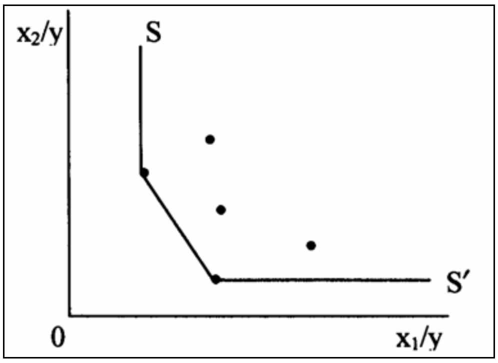

Figure 2: The Data Envelope. Source: Battese and Coelli (1995).

Farrell provided an illustration of his methods using agricultural data for the 48 continental states of the US. The radial nature of Farrell's geometric efficiency measure makes it particularly appealing, given its multiplicative separability property. However, radial measures do encounter consistency problems whenever input reduction has to be

[6] One could illustrate this by drawing two isocost lines through Q and Q'. Irrespective of the slope of these two parallel lines (which is determined by the input price ratio, or relative price), the ratio RQ/OQ would represent the percentage reduction in costs associated with movement from Q to Q'.
[7] Translog functional forms were first introduced by Christensen, Jorgenson, and Lau (1971, 1973). Using Monte Carlo simulations, Guilkey, Lovell, and Sickles (1983) compare three flexible functional forms - translog, generalised Leontief (Diewert), and generalised Cobb-Douglas - and conclude that the translog form performs at least as well as the other two and "[...] provides a dependable approximation to reality, provided that reality is not too complex" (IER, page 614).

measured along segments which are parallel to one of the axes. In those cases, 'input slacks' are found, and alternative, non-radial efficiency measures should be used[8].

If one wishes to know by how much output quantities can be proportionally expanded without altering the input quantities being used, then an 'output-orientated' measure must be introduced. The difference between the output and input-orientated measures can be illustrated by using a simple example involving one input and one output. This is depicted in Figure 3a, where we have a decreasing returns to scale technology represented by $f(x)$, and an inefficient firm operating at point P. The Farrell input-orientated measure of TE is AB/AP, while the output-oriented measure of TE is CP/CD. The output and input-oriented measures will only provide equivalent technical efficiency values when constant returns to scale (CRS) are assumed, but will be unequal when either increasing or decreasing returns to scale are encountered[9].

The case where production involves two outputs (making up the output vector **Y**), a single input (say, x), and constant returns to scale is depicted in Figure 4, where the line ZZ' is the unit production possibility curve, and point A corresponds to an inefficient DMU. Notice that the inefficient point (A) lies below the curve in this case, because ZZ' represents the upper bound of production possibilities.

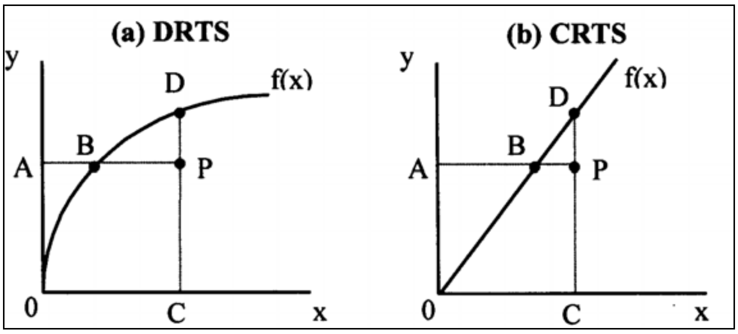

Figure 3 (a, b): Input- and output-orientated efficiency measures. Source: Battese and Coelli (1995).

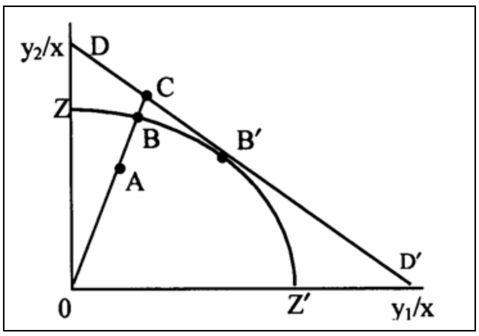

Figure 4: Relative efficiency measurement in output-output space. Source: Battese and Coelli (1995).

[8] See Fare et al. (1985, 1994) on the 'Russell' measure of microeconomic efficiency.
[9] Fare and Lovell (1978).

The Farrell output-oriented efficiency measures are defined as follows. In Figure 4 above, the AB distance represents technical inefficiency (that is, the amount by which outputs could be increased without requiring extra inputs). Hence, a measure of output-oriented technical efficiency will be the following ratio:

TE (*output*) = OA/OB. (5)

If we have price information, then we can draw the iso-revenue line DD', and define allocative efficiency to be equal to

AE (*output*) = OB/OC. (6)

Output-oriented allocative efficiency (AE) has now a revenue-increasing interpretation, which is similar to the cost-reducing interpretation of allocative inefficiency in the input-oriented case. Furthermore, one may define overall output-oriented economic efficiency as the product of TE and AE (again, notice multiplicative separability), that is

$$EE_{output} = \frac{OA}{OC} = \frac{OA}{OB} \cdot \frac{OB}{OC} = TE_{output} \cdot AE_{output}. \qquad (7)$$

Again, all of these three measures are bounded by zero and one. Before we jump to more detailed decompositions of technical efficiency[10] and to the theoretical treatment of DEA[11], two quick points should be made regarding the six efficiency measures which were previously defined:

(*a*) all the above efficiency measures are computed along a ray from the origin of axes to the observed production point. Hence, they hold the relative proportions of inputs (or outputs) constant. One advantage of these radial efficiency measures is that they are invariant to the choice of measurement units. A non-radial measure, such as the shortest distance from the production point to the production surface, may be argued for, but this measure will not be invariant to the units of measurement being chosen. Changing the units of measurement in this case could result in the identification of a different 'nearest' point. This issue will be tackled when we come to consider the treatment of 'input slacks' (or 'excesses') in DEA;

(*b*) the Farrell input and output-oriented technical efficiency measures may be shown to be equal to the input and output distance functions discussed in Shephard[12] (1970). This observation becomes important when one considers the use of DEA methods in calculating Malmquist indices of Total Factor Productivity (TFP) change[13].

Fare, Grosskopf, and Logan (1983) analysed the efficiency of a sample of Illinois electric utilities. They decomposed global economic efficiency (EE) into three further measures:

$$EE = AE \cdot TE = AE \cdot (SE \cdot CE \cdot PTE), \qquad (8)$$

where *AE* = input-allocative efficiency, *SE* = scale efficiency, *CE* = congestion efficiency, and *PTE* = purely technical efficiency. The calculation of these three measures is illustrated in Figure 5. Technical inefficiency is decomposed into three further multiplicative elements.

---

[10] See Fare et al. (1985).
[11] Charnes, Cooper, and Rhodes (1978) is the seminal paper.
[12] For more on this, see Lovell (1993).
[13] Refer to Forsund and Kittelsen (1997) for Norwegian electricity distribution over the Eighties and Nineties. Burns and Weyman-Jones (1994b) provide an empirical application to regional electricity distributors in England and Wales.

Scale efficiency (*SE*) relates to the fact that the DMU might not operate at optimal scale. If one assumes that an 'ideal' unit isoquant for CRS may be drawn, then the radial distance between such isoquant and the 'real' (VRS) one will give a measure for 'scale inefficiency' - that is, inefficiency stemming from the firm operating on a non-flat region of the AC curve, where returns to scale are either increasing (IRS) or decreasing (DRS). Whereas scale efficiency assumes that the initial CRS assumption of Charnes, Cooper, and Rhodes (1978) is dropped, congestion efficiency (*CE*) may be computed provided that the assumption of strong (i.e., free) disposability of inputs is ignored by the linear program. This allows the unit isoquant to bend backwards, indicating that as the quantity of one input (say, *K*) is increased, the output produced falls. Finally, purely technical (in)efficiency (*PTE*) is residually defined as the measure of how much inefficiency is solely due to merely technical gap between the DMU under consideration and the best-practice DMU, or best-practice 'peer', to which the current unit is being compared. In other words, PTE is 'pure' inefficiency after that both scale (VRS/CRS) and congestion inefficiencies have been extracted from the global *TE* (technical efficiency) measure. All *SE*, *CE*, and *PTE* are bounded by zero and one, and multiplicatively recover the *TE* measure, so that *PTE* may be computed residually, after knowing *TE*, *CE*, and *SE*.

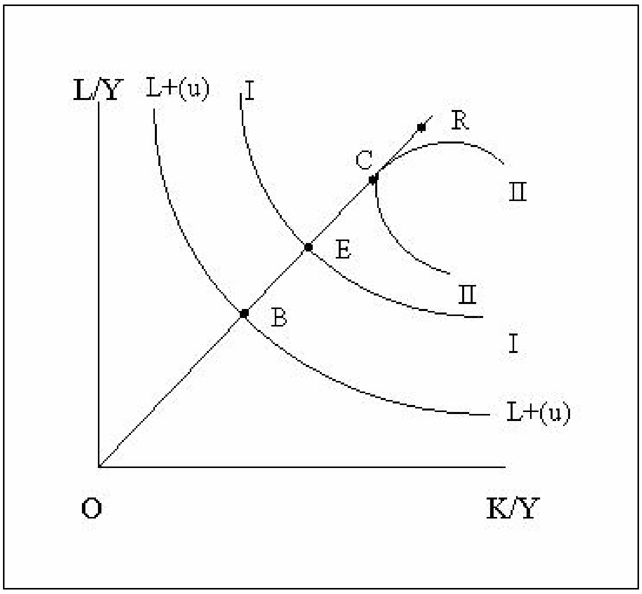

Figure 5: Decomposition of overall (technical) inefficiency into scale, congestion, and 'purely technical' components. Source: Pollitt (1995).

In Figure 5 - drawn in (*K*, *L*) unit space - the above efficiency measures are

SE = OB/OE; CE = OE/OC; PTE = OC/OR, so that (9)

$$TE = SE \cdot CE \cdot PTE = \frac{OB}{OE} \cdot \frac{OE}{OC} \cdot \frac{OC}{OR} = \frac{OB}{OR} \in (0,1]. \quad (10)$$

Whereas congestion inefficiency originates from the fact that **C** falls on a backward bending section of the **II** isoquant, so that inputs are congesting, scale inefficiency is due to the fact that, in the long run, DMUs in competitive equilibrium should be operating at a minimum of the U-shaped AC curve. Therefore, the DMU under examination in the short run might adjust its scale and produce its current output level with fewer inputs. The 'best-practice' unit isoquant **L+(u)**, where all efficient 'peers' are located, is the relevant isoquant for comparison as it incorporates the twin, 'ideal' assumptions of strong (free) input disposability (leading to CE = 1) and constant returns to scale (leading to SE = 1). It is obvious that, under the above assumptions, both scale and congestion inefficiencies will disappear, thus making TE $\equiv$ PTE, and EE = AE $\cdot$ PTE. The details of the linear programs required for the above three measures of technical efficiency (SE, CE, and PTE) are given in Fare et al. (1985, 1994), and Pollitt (1995). It is also possible to identify whether a DMU showing scale inefficiency is exhibiting constant, increasing, or decreasing returns to scale, and thus classify it as CRS, IRS, or DRS. Modern DEA computer packages are able to compute the above efficiency measures for each DMU in the sample, and also tell whether DMUs are operating on the decreasing (IRS), flat (CRS), or increasing (DRS) region of the U-shaped average cost curve.

Data Envelopment Analysis (DEA) is the non-parametric mathematical programming approach to frontier estimation. The discussion of DEA models being presented here is brief, with relatively little technical detail[14]. The piecewise linear convex hull approach to frontier estimation, proposed by Farrell (1957), was considered by only a handful of authors in the two decades following Farrell's paper. Authors such as Boles (1966) and Afriat (1972) suggested mathematical programming methods which could achieve the task, but their method did not receive wide attention until a paper by Charnes, Cooper, and Rhodes (CCR, 1978) inserted Afriat's methodology within a standard operations research framework, and coined the term 'Data Envelopment Analysis' (DEA). Since then, there have been a large number of papers extending and applying the DEA methodology. CCR proposed a model which had an input orientation and assumed constant returns to scale (CRS). Subsequent papers have considered alternative sets of assumptions, such as Banker, Charnes, and Cooper (BCC, 1984), who proposed a variable returns to scale (VRS) model.

3. Stochastic Frontier Estimation: a Theoretical Summary

Deterministic frontier analyses have been most used in the Sixties and early Seventies[15]. The need for separating efficiency errors - i.e., those due to the firm erroneously shifting within its production possibilities set - from purely random noise led theoretical researchers in production econometrics to devise a brand-new framework which was capable of dealing with 'efficiency errors' (one-sided), as separated from either noise or imperfect information[16].

The stochastic frontier production function was independently proposed by Aigner, Lovell, and Schmidt (1977), and Meeusen and van den Broeck (1977). The original specification involved a production function for cross-sectional analysis, featuring an error term which had two components, the first one to account for random effects (a traditional, two-sided disturbance term), with the second one accounting for technical inefficiency (a one-sided error). This model can be expressed in the following form:

---

[14] More detailed reviews of the methodology are presented by Seiford and Thrall (1990), Lovell (1993), Ali and Seiford (1993), Lovell (1994), Charnes et al. (1994, 1995), Pollitt (1993, 1995), and Seiford (1996). Theoretical perspectives on DEA are also outlined, within a principal-agent relationship, by Bogetoft (1994).
[15] Nerlove (1963), in Zellner (1968); Christensen and Greene (1976).
[16] Hebden (1983); McElroy (1987).

$$
\begin{aligned}
&Y_i = X_i\beta + (v_i + u_i), \qquad i = 1,\ldots,N,\\
&\textit{where}:\\
&Y_i = \text{ production (possibly logged) of the } i\text{ - th firm;}\\
&X_i = \text{ a } k \times 1 \text{ vector of (transformations of the) input quantities of the } i\text{ - th firm;}\\
&\beta = \text{ a vector of unknown parameters;}\\
&v, u = \text{ two separate random variables.}
\end{aligned}
\tag{21}
$$

The two-sided random error (v) is assumed to be identically and independently distributed as a normal, with zero mean and constant variance. In particular, such a traditional two-sided random disturbance is independent of u, which is assumed to be a non-positive random variable accounting for technical inefficiency in production. The 'efficiency error' u is often assumed to have a truncated normal, half-normal, gamma, or exponential distribution[17]. If a cost function is used instead of a production relationship, the one-sided error will be non-negative, thus reflecting efficiency errors leading the firm to shift above its cost-minimising contour.

The original stochastic frontier specification has been used in a vast number of empirical applications over the past two decades. The above, standard specification has also been altered and extended in a number of ways. These extensions include the specification of more general distributional assumptions for the efficiency error (u), such as the two-parameter gamma distribution; the consideration of panel data and time-varying technical efficiencies; the extension of the methodology to cost functions and also to the estimation of equation systems. A number of comprehensive reviews of this literature are available, such as those proposed by Forsund, Lovell, and Schmidt (1980), Schmidt (1986), Bauer (1990), and Greene (1993b).

Going back to the one-sided or 'efficiency' error, it must be emphasised that the (in)efficiency component cannot be observed directly. In fact, it must be inferred from the composite error

$$\varepsilon_i = v_i + u_i. \tag{22}$$

Jondrow, Lovell, Materov, and Schmidt (1982) derived an explicit form that decomposes the total error term. The interested reader is referred to their paper and also to Battese and Coelli (1988, pg. 392-393).

Cost efficiency in SFA is simply given by[18]

$$EFF_i = \exp(u_i) \lesseqgtr 1, \tag{24}$$

depending on the main relationship being either a production or a cost function, respectively. Notice that, if the dependent variable is not expressed in natural logs, the above expression will not be valid. The following one should be used instead:

$$EFF_i(linear) = \frac{X_i\beta + u_i}{X_i\beta} \begin{cases} \in (0,1] \text{ for a stochastic production frontier;} \\ \in [1,+\infty) \text{ for a stochastic total cost frontier.} \end{cases} \tag{25}$$

[17] See also Greene (1990).
[18] This is computed by software packages in order to construct efficiency scores for every firm in the sample, being used to build up a final 'efficiency ranking' of all observed units.

Jondrow et al. (1982) also derived similar expressions for exponentially-distributed efficiency errors, whereas Stevenson (1980) did the same for the truncated model. Moreover, Schmidt and Sickles (1984) identified three shortcomings of the cross-sectional estimation of stochastic frontiers: first, the assumption that firm-specific inefficiency is uncorrelated with the explanatory variables can be violated; secondly, the error term may not always be normally distributed; finally, the estimate of u, the efficiency error, may not be consistent. Panel data estimation of stochastic frontier models, which is able to overcome the above limitations, is reviewed and implemented by Burns and Weyman-Jones (1994a). Battese and Coelli (1988) provided a panel data counterpart to the above reported cross-sectional decomposition of the error term. As it will be clarified later, estimation of both cross-sectional and panel data stochastic frontier models is carried out by Maximum Likelihood Estimation (MLE) techniques.

Battese and Coelli (1992) propose a stochastic frontier production function for (unbalanced) panel data which has firm-specific efficiency effects being distributed as truncated normal random variables. Such firm-specific efficiency errors are also permitted to vary systematically with time. A cost-function alternative to more traditional production functions is also provided by the authors. There are obviously a large number of model choices that could be considered for any particular application. For example, a half-normal distribution for the (in)efficiency effects might be assumed, instead of the more general truncated normal distribution. Furthermore, provided that panel data is available, one could assume either time-invariant or time-varying inefficiencies. One could even revert to deterministic OLS estimation if she believes that the u term is not significant, and should then be removed from the model altogether.

Traditionally, a number of empirical studies[19] have estimated stochastic frontiers and predicted firm-level efficiencies by using these estimated functions. Then, they regressed predicted efficiency scores on environmental variables (including ownership types) in an attempt to identify some of the reasons for differences in predicted efficiency scores among firms in a given industry. This has long been recognised as a useful exercise. Lovell (1993, pg. 53), with reference to both stochastic frontier and DEA analyses, points out that "[...] It is worth thinking hard about what variables are inputs and outputs that belong in the first stage [i.e., either production/cost estimation or first-stage DEA], and what variables are explanatory variables that belong in the second stage. This must be done on a case-by-case basis, of course. The only general guideline I have to offer is that variables under the control of the decision-maker during the time period under consideration belong in the first stage. Variables over which the decision-maker has no control during the time period under consideration belong in the second stage. Candidates for second-stage variables include quasi-fixed variables, site-specific characteristics, socio-economic and demographic characteristics, the weather, and so on". The author also discusses the possibility of estimating efficiency scores by limited-dependent variables techniques (e.g., Tobit) because of the censoring problem resulting from the 0/1 scale.

However, as Coelli (1996a) observes, the two-stage estimation procedure has also been long recognised as one which is inconsistent in its implicit assumption regarding the independence of the inefficiency effects (the us) in the two estimation stages. The two-stage estimation procedure is unlikely to provide estimates which are as efficient as those that could be obtained by using a single-stage estimation procedure. This issue was addressed by Kumbhakar, Ghosh, and McGuckin (1991), and Reifschneider and Stevenson (1991), who proposed stochastic frontier models in which the inefficiency effects (u) are expressed as an explicit function of a vector of firm-specific variables plus a random error.

[19] E.g., see Pitt and Lee (1981).

4. A Comparison Between DEA and Stochastic Frontier Estimation

In this chapter we sketch the main pros and cons which should be taken into consideration when comparing DEA efficiency results with those stemming from stochastic frontier analysis, an econometric technique often used by regulators as a cross-check on DEA[20]. In order for methodological cross-checking to be fully understood, the following differences between DEA and Stochastic Frontier Estimation (SFE) must be kept in mind:

(*a*) the DEA technique does not require the specification of a functional form for the production function. Such flexible functional forms as the translog definitely improve the situation, but are not able to clarify how different components of efficiency might be separated within SFE. As noticed in the previous Sections, stochastic frontier methods are generally unable to effectively distinguish between input-allocative and technical efficiency. Battese and Coelli's (1995) 'Technical Efficiency Effects' model, for instance, sorts out the problem by simply assuming that allocative efficiency is *ex ante* full (AE $\equiv$ 1), thus ascribing all deviations from best-practice efficiency to technical effects (apart from random noise, of course). In general, SFE encounters serious problems with input-allocative efficiency computations;

(*b*) the DEA technique is non-stochastic[21] and does not take errors (random noise, measurement error) into consideration, unlike the SFE technique. Empirical work always involves some degree of measurement error, data handling errors, stochastic shocks, *et similia*. On the other hand, the ability of SFE to handle errors only comes at the expense of either the imposition of a functional form for the errors themselves[22], or the use of panel data[23];

(*c*) DEA scores are heavily affected by specification and variable selection errors. On the contrary, SFE frontiers give rise to standard errors and *t*-values for each of the parameters (e.g., in a translog total cost function). It is also true that sophisticated sensitivity analysis of DEA scores might be a possible solution to this problem, but it would be indeed complex and time-consuming;

(*d*) from a computational point of view, differences between DEA and SFE are now negligible, as a consequence of modern computing power. SFE involves econometric estimation of either a production or a cost function (plus share equations, if feasible), whereas DEA implies a separate linear program for each DMU in the sample. The most recent DEA software is capable of accommodating thousands of simultaneous linear problems automatically, so that computer programming skills are no more essential;

(*e*) DEA allows easy extension to multiple outputs in a production frontier context, and can also accommodate non-discretionary variables into the analysis in a direct way. On the contrary, SFE may only be extended to multiple outputs within a cost frontier setting. However, an excessive number of outputs (or inputs) in DEA will rapidly give rise to an excessive number of best-practice DMUs. This may result in a tighter constraint on the

[20] The Austrian regulator E-Control has recently announced that both DEA and stochastic frontiers will be used in its comparative efficiency exercise for price regulation purposes, but the econometric technique will only be used as a cross-check.

[21] Some researchers have tried to develop a stochastic version of DEA (see Land, Lovell, and Thore, 1988, and Lovell, 1993), but those models are by no means complete, and require massive data information regarding expected values, variance-covariance matrices, probability levels, and so on. Stochastic DEA would solve the main methodological problem with Data Envelopment Analysis, and is therefore worth investigating to a larger extent in the future.

[22] See Jondrow, Lovell, Materov, and Schmidt (1982), and Greene (1993b).

[23] See Burns and Weyman-Jones (1994a).

number of inputs/outputs which can be inserted in DEA - as opposed to SFE - without saturating the model[24].

5. Real-world regulatory usage of economic benchmarking: lessons from the Dutch experience with electricity networks' benchmarking

The Dutch energy regulator DTe is responsible for setting tariffs for electricity transmission and distribution networks. This is done on the basis of a price cap system according to a CPI-X formula. DTe published its first decision on the X factors for electricity transmission and distribution networks in September 2000. These X factors were strongly driven by the results of a DEA benchmark report. Three separate DEA benchmarks were carried out: For the national transmission grid (TenneT), for a regional transmission grid (TZH), and for the regional network companies. The DEA benchmarks for TenneT and TZH were later discarded by DTe. In the remainder of this paper, we will focus on the DEA benchmark for the regional networks.

Input factor	Total Costs [NLG]
Output factors	Energy delivered [kWh] Number of small customers (distribution) Number of large customers (transmission) Peak demand at distribution level [MW] Peak demand at transmission level [MW]
Environmental factors	Number of transformers Network route length [km]

Table 1: DEA benchmark model specification (DTe, the Netherlands, 2000).

The DEA benchmark was applied to the sample of 20 Dutch local network companies. The model specification is shown in table 1. DTe used a single input factor, which was the total cost of each network company. This total cost was the sum of operating expenditure, depreciation, and a standard return on assets. Certain cost elements, which were considered to be uncontrollable by the network company, were not included in the benchmark analysis and were remunerated separately on the basis of actual costs. A peculiar feature of DTe's model is its total cost approach.

A special feature of the Dutch regulation system is that both operating and capital costs were considered in the DEA efficiency analysis (DTE 2002). Operating expenditure is associated with personnel, maintenance etc. while capital expenditure is made up of depreciation plus a fair rate of return on investment. As far as known to the authors, no other regulator directly includes capital costs in the efficiency analysis. In other regulatory cases, the calculation of the X-factor also includes capital costs but these are not directly benchmarked. Benchmarking analysis is restricted to operating costs whilst (scrutinised) capital costs are simply added on a cost-plus basis when setting the tariffs. Traditionally, DEA benchmarks are carried out either for operating expenditure alone, or when capital expenditure is included, this is usually through the use of proxy parameters such as the number of installed transformers, network length, etc.

The DTe approach is different in that cost data, rather than technical inputs, were used and that operating and capital costs were considered simultaneously. DTe's motivation for this approach was that it is the company's responsibility to trade-off between short and long-term costs. Insofar as this trade-off had any effect on the overall efficiency of the company, it should be reflected in the efficiency analysis. By simultaneously considering operating and capital expenditure in the efficiency analysis, DTe bypassed the regulatory problem of investment appraisal. Experience shows that this is one of the main difficulties

[24] On the 'parsimony' requirement for DEA, see Ferrier and Lovell (1990).

faced by regulators. Under a total cost model, this problem was simply bypassed as there was no explicit requirement anymore to consider CAPEX projections in the X factor calculations. In principle, a total cost approach is therefore preferable because it creates incentives to improve efficiency in the short as well as the long term. Given that capital forms a significant portion of total cost, the effectiveness of regulation can be greatly enhanced. However, benchmarking capital costs is extremely difficult due to data problems. Capital costs reflect the investment process and exhibit long- term characteristics that imply multi-period determinations of depreciation and of the return on assets. For instance, different companies may use different depreciation profiles (asset valuation, asset life, depreciation path) as allowed by national accounting rules. DTe aimed to eliminate such monetary effects resulting from bookkeeping practice by performing a backward calculation of book and depreciation values. In doing so, however, a number of assumptions and approximations had to be made. Due to the lack of detailed data, the standardisation was performed on an aggregate basis, thereby ignoring the differences in lifetime and age across asset categories. Also, as historical investment profiles were not available, a virtual annual investment profile was assumed when recalculating the asset and depreciation values.

DTe used three types of outputs, namely energy delivered, number of customers, and peak load. For the latter two outputs, a distinction was made between distribution and transmission. While some companies had only distribution activities, others had a mix of transmission and distribution. To accommodate these differences, the data on customers and peak load was split between distribution and transmission. In addition to the total of five output factors, DTe included two environmental factors in the benchmark. Network route length was used as a proxy for the size of the network, while the number of transformers acted as a proxy for complexity.

The September 2000 decisions on the X factors led to a wave of protest and formal appeals by the industry. Their main critique was aimed at the use of benchmarking as a way to set tariffs, the mechanical translation of efficiency scores into X factors, and the use of limited and flawed data – in particular, the standardisation of capital costs. Additionally, the fact that DTe widely published the benchmarking results did not help in this regard. As a result, the relationship between regulator and industry became increasingly hostile: on the one hand, DTe confirmed its decisions; on the other hand, the network companies refused to accept the – in their eyes unjust and erroneous – X factor decisions. The appeals led DTe to publish revised decisions in September 2001. The main difference with the initial decisions was an increase in the quality of data. An independent audit was performed to verify and improve the output factor data, while the CAPEX standardisation was refined by considering each individual asset and the actual historical investment profile. The data improvement led to higher efficiency scores and lower X factors. However, the companies' main critique points were still not thoroughly met, and there still remained problems with the data. DTe responded to this by initiating a special project with the objective to remove any remaining data problems. As a result, a second revision of the benchmark analysis and X factors was published in August 2002. But this did not prevent the network companies from confirming their appeals, as they did not consider DTe's corrections to be sufficient. Eventually, in October 2002, the Courts overruled the X factor decisions. However, this was just a legally-motivated decision. It was taken not (only) on the basis of the benchmark analysis, but mainly because according to the Dutch Electricity Act, DTe should have applied a uniform X factor (instead of an individual X factor for each company) in the first place.

6. Benchmarking price and quality

In this Section, we present a benchmark model based on Data Envelopment Analysis (DEA), which we use to carry out integrated analysis of cost and quality for distribution firms. For quality, we use annual minutes "not served" (Customer Minutes Lost or CML). We make use of a multi-national cross-sectional sample of distribution operators from the Netherlands, the UK, Hungary, and Malaysia. The data was gathered from publicly available sources including regulatory publications and company accounts. We analyse the sensitivity of different parameters and perform crosschecks with other, statistically based benchmarking techniques to test the significance and validity of our linear programming (DEA) outcomes.

Data Envelopment Analysis (DEA) takes, as its starting point, that - for any production process - there is an output frontier linking the maximum combination of outputs attainable with a given set of inputs, or, equivalently, for any set of outputs, there is an input frontier linking minimum combinations of inputs that are needed to produce those outputs. Production units that lie on this frontier are said to be "technically efficient". Technical efficiency is measured by comparing each production unit's actual combination of inputs and outputs with a relevant point on the frontier. For cost-minimising DEA, input price information would be required. In the absence of reliable information on input prices, we will here use a hybrid form of DEA whereas cost is treated as a - physical-like - input, and therefore the program is run as a technical efficiency DEA. We are aware of the theoretical problems caused by the usage of cost as a minimisable "input" in DEA, and we refer the reader to DTe's paper on cost-based benchmarking for more information on this hybrid DEA methodology and the theoretical justifications for it (DTe, 2000).

There are different ways of carrying out cost-based DEA, and these may lead to different results. We opted for a cost minimising technical efficiency model, because of the unavailability of data on input prices and thus the practical difficulty of separating out technical and "allocative" inefficiency. As regards scale assumptions in DEA, it is good to include the option between "constant" and "variable" returns to scale. As opposed to econometric techniques, DEA's application is easier in relating efficiency scores to non-economic features. This is mainly due to the fact that DEA computes scores rather than estimating them, and that no functional form is imposed upon technology. The absence of restrictions entails the "encompass" effect typical of DEA, leading to the absolute reliance of this technique on any single data point available, without any statistical smoothing. The use of DEA, therefore, makes quality analysis possible once quality is quantified in a reliable way, without the imposition of any functional form at all on the cost/quality relationship.

A standard DEA specification is the one considering total expenditure (TOTEX) as an input, both without (Model 1) and with quality (Model 2). The Model 2 variation on the basic DEA specification (Model 1) is the one using Customer Minutes Lost (CML) as a 'negative output', i.e. a fictitious input to minimise. The CML model is useful as a sensitivity variation on the standard totex specification. In all cases, we have two conventional distribution DEA outputs (to be held fixed in input- minimisation mode), i.e. kWh being distributed and the number of connections (approximated by the number of customers). The two models we envisage for our analysis are reported below in Table 1.

	Model 1	Model 2
TOTEX	Input	Input
CML		"Negative Output" (minimis
KWH	Output	Output
CUSTOMERS	Output	Output

Table 2: DEA model specifications.

TOTEX is calculated as the sum of OPEX and Depreciation[25], at PPP-adjusted (April 2003) exchange rates normalised to the US Dollar (USD) in thousands. CML is defined as the cumulative minutes of interruption for all customers in any given year[26]. Energy delivered and customers are as reported in regulatory accounts by the utilities. All variables are re-scaled in natural (base e) logarithms for the purposes of econometric estimation, whereas they enter DEA as natural values.

Companies included in the dataset come from the UK, the Netherlands, Hungary (all of which are EU member states in 2004), and Malaysia from the ASEAN group. The data are for 2003 and the cross-section is internationally pooled. It excludes one UK utility for which depreciation data were not computable – thus bringing the number of UK distribution firms (DNOs) to thirteen from the fourteen making up mainland Britain (Northern Ireland is subject to a different industry and regulatory structure, and is therefore excluded). Otherwise, the sample is made up of six regional Hungarian utilities (distribution only), twelve Dutch territorial distributors (this number continuously decreasing due to merger activity), and thirteen Malaysian federate states from the territorial split of Tenaga Distribution, the distribution arm of Malaysia's lead electric utility TNB.

DEA models are run under both a constant and variable returns to scale assumption (CRS and VRS, respectively, and Models 1 and 2). This is to accommodate the debated issue of scale efficiency and responsibility for scale optimisation on the utilities' side, an issue that contributed to cripple the first DTe (NL) electricity distribution benchmarking exercise of 1999-2000.

Econometric cross-checking models are run in natural-logarithm mode taking cost as the endogenous (dependent) variable, and all other variables as explanatory (exogenous) factors. The functional form chosen for the econometric cross checks is the standard, log-log Cobb-Douglas cost function that is, on most occasions, a valid nested approximation to, and an acceptable statistical restriction of the more general, but much more data demanding, "Translog" cost function. A nice corollary of this choice is that the estimated parameters from the cost function can directly be interpreted as (cost-output) elasticities.

DEA and econometric results (efficiency scores) are readily comparable on a statistical basis by means of both parametric (t) and non-parametric (distribution-free) tests. Efficiency scores are censored at 100% and cannot be lower than 0%. Under these conditions the usual testing procedure for statistical comparisons across score series would have to be non-parametric, as score series will sometimes fail to be normally distributed. However, normality can be tested in advance in order to select the preferred statistical cross-checking methodology.

The models we ran, Models 1 and 2, gave DEA and COLS (Corrected Ordinary Least Squares) scores to cross-check. The total sample of 44 cross-sectional observations is too small to guarantee acceptable results in the presence of more complex statistical assumptions such as those underlying Stochastic Frontier Analysis (SFA). We tried to calculate SFA scores, but these were all clustered towards full efficiency, thus losing any discriminatory power in terms of efficiency ranking. This often happens to SFA with relatively small samples. The uselessness of SFA results led us to drop this technique from the modelling exercise.

On the other hand, COLS is a relatively accepted and unambiguous modelling technique to cross-examine DEA results. It is based on conventional OLS estimation, which is corrected

[25] We ignore here the return on assets (equal to RAB times WACC) for reasons of lack of comparability, in particular for non-EU countries.
[26] In our analysis, CML was calculated as the product of SAIDI (System Average Interruption Duration Index) and the total number of customers.

by moving the regression line either back or forward by a horizontal shift in order to encompass the whole dataset. By so doing, COLS heavily relies on the position of the single most efficient observation in the dataset, and is in a sense similar to DEA in terms of extreme data dependence. However, COLS and DEA efficiency scores are calculated on the basis of completely different assumptions otherwise, and therefore they do not easily agree with each other in statistical terms. This introduces a degree of uncertainty in the benchmarking exercise if the latter is used by regulatory authorities for price control purposes.

Efficiency scores (bounded by zero and one) from both Models 1 (no quality) and 2 (CML) follow for all of the (anonymous) 44 utilities making up our cross section.

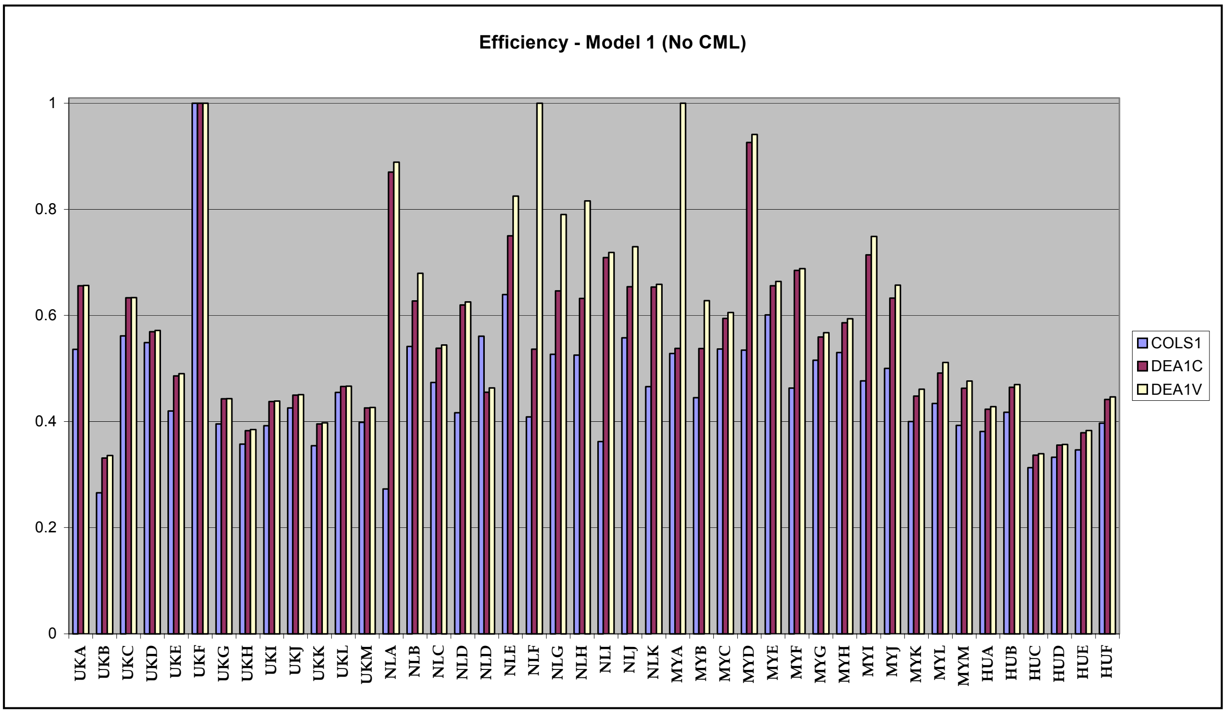

Figure 6: Efficiency Scores without Quality.

	Company	COLS1	DEA1C	DEA1V	COLS2	DEA2C	DEA2V
1	UKA	0.54	0.66	0.66	0.53	0.85	1.00
2	UKB	0.27	0.33	0.34	0.30	0.39	0.40
3	UKC	0.56	0.63	0.63	0.58	0.71	0.80
4	UKD	0.55	0.57	0.57	0.57	0.66	0.70
5	UKE	0.42	0.49	0.49	0.44	0.59	0.61
6	UKF	1.00	1.00	1.00	1.00	1.00	1.00
7	UKG	0.40	0.44	0.44	0.40	0.55	0.69
8	UKH	0.36	0.38	0.38	0.37	0.49	0.54
9	UKI	0.39	0.44	0.44	0.41	0.49	0.52
10	UKJ	0.43	0.45	0.45	0.43	0.59	0.78
11	UKK	0.35	0.40	0.40	0.36	0.54	0.65
12	UKL	0.45	0.47	0.47	0.46	0.56	0.62
13	UKM	0.40	0.43	0.43	0.40	0.58	0.82
14	NLA	0.27	0.87	0.89	0.30	1.00	1.00
15	NLB	0.54	0.63	0.68	0.54	0.91	1.00
16	NLC	0.47	0.54	0.54	0.46	0.78	0.92
17	NLD	0.42	0.62	0.63	0.41	0.80	0.88
18	NLD	0.56	0.46	0.46	0.57	0.65	0.80
19	NLE	0.64	0.75	0.82	0.70	0.86	0.88
20	NLF	0.41	0.54	1.00	0.46	0.70	1.00
21	NLG	0.53	0.65	0.79	0.49	1.00	1.00
22	NLH	0.53	0.63	0.82	0.45	1.00	1.00
23	NLI	0.36	0.71	0.72	0.36	0.91	1.00
24	NLJ	0.56	0.65	0.73	0.56	0.95	1.00
25	NLK	0.47	0.65	0.66	0.42	0.93	1.00
26	MYA	0.53	0.54	1.00	0.53	0.61	1.00
27	MYB	0.45	0.54	0.63	0.45	0.77	0.78
28	MYC	0.54	0.59	0.61	0.56	0.68	0.71
29	MYD	0.53	0.93	0.94	0.59	1.00	1.00
30	MYE	0.60	0.66	0.66	0.63	0.68	0.68
31	MYF	0.46	0.68	0.69	0.51	0.73	0.77
32	MYG	0.52	0.56	0.57	0.55	0.64	0.66
33	MYH	0.53	0.59	0.59	0.58	0.59	0.59
34	MYI	0.48	0.71	0.75	0.54	0.79	0.80
35	MYJ	0.50	0.63	0.66	0.54	0.76	0.77
36	MYK	0.40	0.45	0.46	0.44	0.48	0.48
37	MYL	0.43	0.49	0.51	0.47	0.57	0.58
38	MYM	0.39	0.46	0.48	0.38	0.58	0.63
39	HUA	0.38	0.42	0.43	0.41	0.42	0.43
40	HUB	0.42	0.46	0.47	0.45	0.46	0.47
41	HUC	0.31	0.34	0.34	0.34	0.37	0.39
42	HUD	0.33	0.36	0.36	0.35	0.40	0.41
43	HUE	0.35	0.38	0.38	0.38	0.38	0.38
44	HUF	0.40	0.44	0.45	0.42	0.44	0.45

Table 3: Efficiency Scores (all models).

Regarding DEA, our Models are split according to the assumption made on scale efficiency (CRS or Constant Returns to Scale for absolute scale efficiency, similarly to what hypothesised by the Dutch regulator DTe in its actual benchmarking exercises, and VRS for Variable Returns to Scale assuming that firms may not be necessarily operating at optimal scale at the time their overall cost efficiency is assessed). In the absence of input prices in input-minimising DEA, efficiency scores would in theory only reflect 'technical' efficiency, not 'allocative' efficiency. This means that any inefficiency deriving from non-optimal input mixes would not be calculated by a technical DEA exercise. However, in our case we are using total expenditure as the overall DEA "input", and therefore our 'technical' DEA problem turns into a hybrid whereby efficiency scores may reflect both technical and allocative inefficiency, with no possibility whatsoever to separate out the two effects due to the practical unavailability of any reliable information on the relative price of production factors.

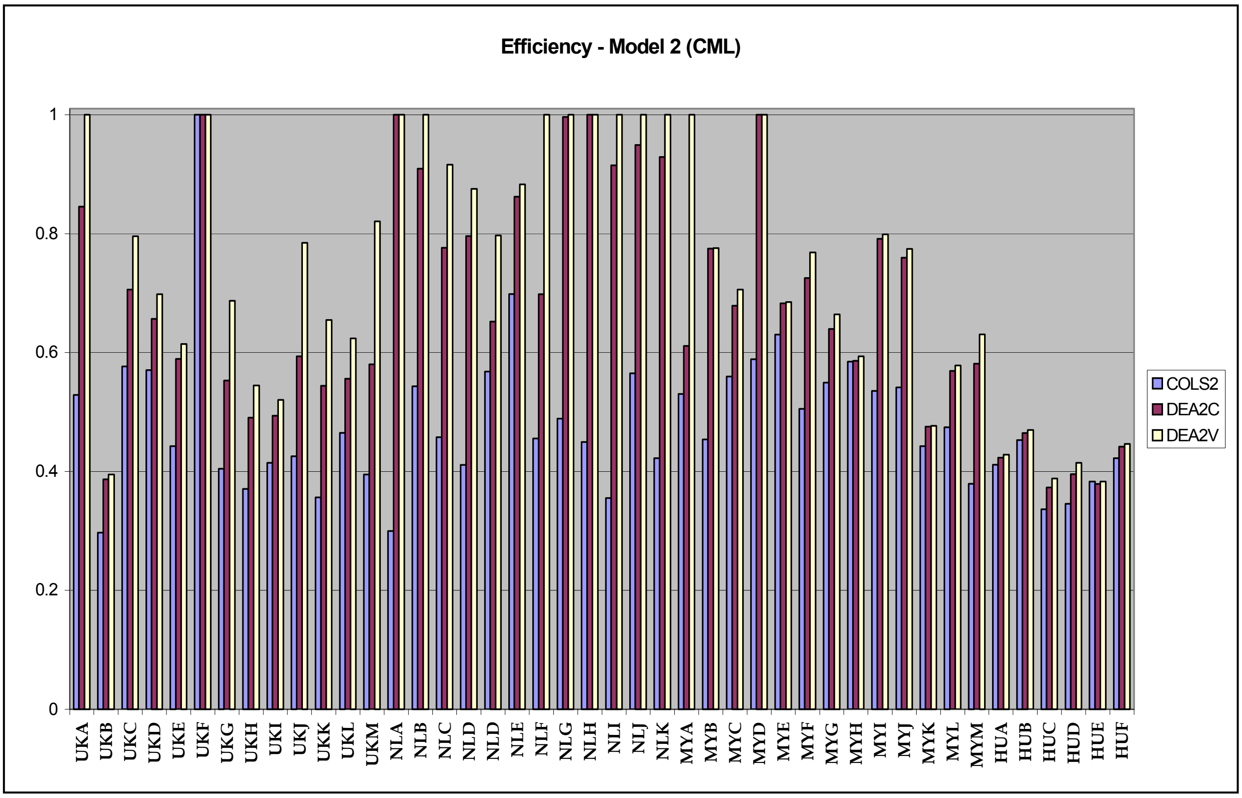

Figure 7: Efficiency Scores with Quality.

In Model 2, average DEA efficiency scores increase. This is a normal consequence of the insertion of extra variables in the DEA program. By cross-checking DEA models against COLS ones, we noticed that the insertion of CML in the cost equation was statistically sensible with a 10%-significant CML coefficient. However, the relative significance of CML in the TOTEX equation somewhat obscured the effect of electricity delivered (KWH) and led us to believe that the 'true' model might only contain customers and CML as the actual determinants of TOTEX.

In Model 2, the presence of CML as a benchmarking variable greatly enhances the relative efficiency of Dutch utilities, especially the smaller ones operating in 'easier' environments. These utilities are generally more cost-efficient than the UK ones on a TOTEX basis, especially when CML is part of the modelling. Generally speaking, the winners from the relative comparisons are the smaller utilities from the Netherlands, even under CRS DEA, and especially in Model 2 once quality of supply is taken into account. With the exception of a notoriously efficient big UK utility, the Dutch ones are relatively better positioned. This may be related to the higher customer density levels in the Netherlands leading to the design of mostly underground and highly meshed networks. In turn, this is associated with fewer outages and shorter outage duration leading to relatively better quality performance.

Interestingly enough, Malaysian and Hungarian distributors do not react to the insertion of CML to the same extent as the British and Dutch firms do. This suggests that British and Dutch firms have been more concerned with producing high quality than their Malaysian and Hungarian counterparts. We may speculate on a number of possible explanations for this. One is that British and Dutch customers have historically demanded higher reliability and this has resulted in higher (regulatory) pressure to deliver high quality. Another explanation may be the relative lower capital costs and associated higher level of investments in these countries. Reliability, which is strongly capital-related, is likely to improve at higher investment levels[27]. However, these are only speculations and further research is needed on this issue.

Another interesting aspect of the results is that Dutch companies gain more than anyone else from the VRS DEA assumption, especially in Model 2. This is due to the fact that many of the Netherlands' utilities are locally owned and based, and that some of them are extremely small, and enjoy a favourable customer density and mix (although this is now changing because of merger activity). CRS DEA will then be harsher to "limit-size" companies, for instance very small ones, which would otherwise obtain much nicer results under a VRS assumption. Interestingly enough, VRS DEA was rejected by the Dutch regulator in 1999 on the grounds that all firms are in fact free to determine their own scale by either merging or de-merging.

A final word should be spent about Hungarian and Malaysian firms. Whereas the latter positively react to the introduction of a quality variable in the benchmarking and become more efficient on average (thus highlighting a cost-quality relationship), the former only marginally improve in Model 2 as opposed to cost-only Model 1. This might be due to the fact that, whereas Peninsular Malaysia has a form of public active quality control and monitoring enacted by the Regulator, the development of quality standards in Hungary is now only beginning, although the Regulator is determined to introduce incentivised quality any time soon. The unimpressive quality performance of Hungarian utilities, in spite of private ownership, can only be tackled by means of strong quality regulation, especially in the presence of regulatory caps imposed on prices and a series of private monopolists
The Peninsular Malaysian situation is institutionally different from Hungary's in that quality performance levels are monitored and published by the Regulator. Moreover, Malaysian local utilities are nothing else than divisions of the state utility TNB, and as such their quality performance is centrally monitored. This sometimes helps quality standards be kept at acceptable levels, albeit at the price of possible gold plating.

Finally, performing statistical testing on matched pairs of COLS and DEA score series highlights a fundamental average difference in the efficiency scores. This outcome is not unusual, and it just confirms that the assumptions on which the statistical and DEA analyses are based differ to such an extent that reaching consistent results is not always viable. For regulatory purposes, the statistical divergence among different efficiency score

[27] For example, in Italy reliability levels have improved strongly mainly as a result of intensified investments in the network (Ajodhia, 2003).

series – even in the presence of the same data set and the same basic model specification – cannot be understated, and may sometimes cast a shadow upon the actual practical applicability of alternative benchmarking techniques for straight regulatory usage.

7. Conclusions

The informational asymmetry problem between the regulator and its regulated agents has been widely discussed in the literature. Lack of information about the network companies' efficiency potential prevents the regulator from setting optimal prices and performance standards. Benchmarking is a powerful tool to reduce this informational asymmetry. In this paper, we have discussed two popular benchmarking techniques.

Data Envelopment Analysis (DEA) is a linear programming technique and calculates efficiency instead of estimating it. This exposes the technique to a host of problems and difficulties relating to the absence of a stochastic component that differentiates between genuine inefficiency and random noise. On the other hand, DEA makes life (apparently) easy to monopoly regulators.

Stochastic Frontier Estimation (or Analysis) is a statistically based technique that attempts at drawing a line between inefficiency and random effects. It does so at the expense of imposing a functional structure upon the production or cost process. This introduces a number of complications in the analysis, and may result in the diminished appeal of this technique for real-word applications. However, regulators are now tending to re-use statistical techniques to cross-check DEA results. This has recently been observed, for instance, in Austria and in a number of Central and Eastern European countries.

Benchmarking techniques should be used cautiously, and their limitations should be recognised. Benchmark studies do not generate the absolute truth, but rather an indication and ranking of relative efficiency levels. Its results will depend strongly on the choice of model, model specification, and the quality of the data. From a regulated company's point of view, it is very difficult to accept the results of an efficiency analysis if changes in data or model specification result into totally different results. In the Dutch case, the results from the benchmarking analysis were directly translated into X factors. This 'linked' approach made the regulatory process extremely sensitive to data errors. A change in the data for a single company might have potentially led to a change in efficiency scores for the whole sample. This also created an opportunity for companies to strategically influence the regulatory process, as each subsequent correction in the regulatory decisions will reduce overall credibility.

The future regulatory regime in Germany will have a significant performance impact on gas and electricity utilities. It will affect the revenue and profits of network service providers as well as the quality of supply. Internationally, incentive regulation has proved to be successful. In Germany, too, RegTP currently applies cap-based regulation supported by incentive instruments to encourage efficiency increases in the telecoms industry. The introduction of incentive-type regulation for German gas and electricity network service providers, supplemented by benchmarking endeavours, should be realistically expected.

The main lesson from the Dutch experience is perhaps that regulator should take into account the limitations of benchmarking. Efficiency scores should be considered as an indication, rather than a confirmation, of (in)efficiency. Given the large degree of uncertainty in the results of a benchmarking study, its outcomes should be used with a grain of salt. Rather than feeding directly into the X factor process, they should be considered as precious pieces of information to strengthen the regulator's position in the discussions with network operators. Benchmarking results are not the solution, but merely the starting point in sorting out the information asymmetry problem.

An appropriate regulatory approach would provide the regulated firm with such economic incentives as to constrain the firm's private profit-maximising strategy towards an outcome that is compatible with socially optimal quality. Complying with this "incentive compatibility" constraint is, however, difficult in practice – in particular, when the quality dimension is included. There are three main informational problems at work. They are discussed in the paper and are briefly reproduced to sum up. Firstly, there is the problem of quality measurement. Secondly, consumer demand for quality needs taking into account. Thirdly, the relationship or 'trade-off' between cost and quality needs proper consideration. Starting from these three problems, we have argued that the solution to the latter is the most difficult to obtain. Until now, integrated analysis of cost and quality in electricity distribution has been very limited. Indeed, the results of such an analysis have not (yet) been formally used and applied for regulatory purposes.

In order to test the third proposition, we performed DEA and COLS analysis on a sample of PPP-adjusted international utility cost and technical/quality variables for 2003, summing up to 44 observations from the UK, the Netherlands, Hungary, and Malaysia. Our findings confirmed that the utilities that are more cost-efficient tend also to better perform on the quality front. The cost-quality trade-off is a dynamic one, so it cannot be captured by cross-sectional analysis. On the contrary, what is more likely to happen in a cross-sectional setting is that the best-efficient utilities will still be best practice once quality comes into play. Moreover, quality-oriented utilities will take the lead in the rankings once a quality output kicks in (the Dutch case), and especially when scale need not influence cost efficiency (VRS DEA). Finally, Malaysian utilities perform better that Hungarian ones in the presence of stronger quality reporting requirements by the Regulator, which makes TNB somewhat closer to Western Europe in this respect than to an accession country such as Hungary. For all the four countries examined, however, higher cost efficiency generally means better quality performances too, which is a consistent result in the presence of a 'snapshot' (cross-sectional) sample. However, when the time dimension comes into play, results may be likely to change and, over a range of – say – five to ten years, the cost-quality trade-off might start to become visible. This is, however, matter for further data research and efficiency analysis.

References

Acharyya, R (1998) 'Monopoly and product quality: separating or pooling menu?' Economic Letters 61 187-94.

Afriat, S.N. (1972). Efficiency Estimation of Production Functions. International Economic Review 13: 568-598.

Ai and Sappington, D (1998) 'The Impact of State Incentive Regulation on the U.S. Telecommunications Industry' University of Florida, December.

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.

Ajodhia, V S (2003) 'Evaluation of the Italian Incentive Scheme for Continuity of Supply' mimeo, Delft University of Technology.

Ajodhia, V S, Petrov, K and Scarsi, G C (2003) 'Benchmarking and its Applications'. Zeitschrift für Energiewirtschaft 27 (4), 261-274.

Ajodhia, V S, Petrov, K and Scarsi, G C (2004) 'Quality, Regulation, and Benchmarking: An Application to Electricity Distribution Networks'. Zeitschrift für Energiewirtschaft 29 (2), 107-120.

Alchian, A.A. (1965). Some Economics of Property Rights. Il Politico 30 (Dec.): 816-829.

Ali, A.I., and L.M. Seiford (1993). The Mathematical Programming Approach to Efficiency Analysis. In Fried, H.O., C.A.K. Lovell, and S.S. Schmidt (eds.) (1993). The Measurement of Productive Efficiency. New York: Oxford University Press.

Andersson, I and Taylor, L (1986) 'The social cost of unsupplied energy' Energy Economics 8 139-46.

Averch, H and Johnson, L (1962) 'Behavior of the Firm under Regulatory Constraint' American Economic Review 52 (5) 1052-69.

Banker, R.D., A. Charnes, and W.W. Cooper (1984). Models for Estimation of Technical and Scale Inefficiencies in Data Envelopment Analysis. Management Science 30: 1078-1092.

Banker, R., R.F. Conrad, and R.P. Strauss (1986). A Comparative Application of Data Envelopment Analysis and Translog Methods: an Illustrative Study of Hospital Production. Management Science 32: 30-44.

Banker, R.D., S. Das, and S.M. Datar (1989). Analysis of Cost Variances for Management Control in Hospitals. In J.L. Chan (ed.): Research in Governmental and Non-Profit Accounting. Greenwich, Conn.: JAI Press.

Banker, R.D., and R.C. Morey (1986). Efficiency Analysis for Exogenously Fixed Inputs and Outputs. Operations Research 34: 513-521.

Battese, G.E., and T.J. Coelli (1988). Prediction of Firm-Level Technical Efficiencies with a Generalised Frontier Production Function and Panel Data. Journal of Econometrics 38: 387-399.

Battese, G.E., and T.J. Coelli (1992). Frontier Production Functions, Technical Efficiency and Panel Data, with an Application to Paddy Farmers in India. Journal of Productivity Analysis 3: 153-169.

Battese, G.E., and T.J. Coelli (1995). A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data. Empirical Economics 20: 325-332.

Bauer, P.W. (1990). Recent Developments in the Econometric Estimation of Frontiers. Journal of Econometrics 46: 39-56.

Baumol, W J and Klevorick, A K (1970) 'Input Choices and Rate of Return Regulation: An Overview of the Discussion' Bell Journal of Economics 1 (1) 162-90.

Bental, B and Ravid, A (1982) 'A Simple Method for Evaluating the Marginal Cost of Unsupplied Energy' Bell Journal of Economics 13 (1) 249-53.

Berg, S.A., F.R. Forsund, and E.S. Jansen (1991). Technical Efficiency of Norwegian Banks: the Non-Parametric Approach to Efficiency Measurement. Journal of Productivity Analysis 2: 127-142.Bericht des Deutschen Budesministeriums für Wirtschaft und Arbeit an den Deutschen Bundestag über die energiewirtschaftlichen und wettbewerblichen Wirkungen der Verbändvreibarhung (Monitoring Bericht), August 2003.

Besanko, D, Donnenfeld, S, and White, L (1987) 'Monopoly and quality distortion: Effects and Remedies' Quarterly Journal of Economics 102 743-68.

Besanko, D, Donnenfeld, S, and White, L (1988) 'The Multiproduct Firm, Quality Choice, and Regulation' Journal of Industrial Economics 36 411-29.

Billinton, Wacker and Wojczynski (1982) 'Customer Damage Resulting from Electric Service Interruptions' Canadian Electrical Association, R&D project 907U 131, April.

Bogetoft, P. (1994). Incentive-Efficient Production Frontiers: an Agency Perspective on DEA. Management Science 40: 959-968.

Boles, J.N. (1966). Efficiency Squared - Efficient Computation of Efficiency Indexes. Proceedings of the 39th Annual Meeting of the Western Farm Economic Association: 137-142.

Boltz, Dr. W., E-Control GmbH, Neue Netztarife, 27. Juni 2003

Bundesverband der Deutschen Industrie e. V. (BDI) – VIK – Verband der industriellen Energie- und Kraftwirtschaftschaft e. V. (VIK) – Verband der Elektrizitätswirtschaft e. V. (VDEW) – Verband der Netzbetreiber (VDN), beim VDEW e. V. – Arbeitsgemeinschaft regionaler Energieversorgungs-Unternehmen e. V. (ARE) – Verband kommunaler Unternehmen (VKU) e. V. 'Verbändevereinbarung über Kriterien zur Bestimmung von Netznutzungsentgelten für elektrische Energie und über Prinzipien der Netznutzung', Essen 2001, 13. Dezember 2001.

Bundesverband der Deutschen Industrie e. V. (BDI) – VIK – Verband der industriellen Energie- und Kraftwirtschaftschaft e. V. (VIK) – Verband der Elektrizitätswirtschaft e. V. (VDEW) – Verband der Netzbetreiber (VDN), beim VDEW e. V. – Arbeitsgemeinschaft regionaler Energieversorgungs-Unternehmen e. V. (ARE) – Verband kommunaler Unternehmen (VKU) e. V.: 'Verbändevereinbarung über Kriterien zur Bestimmung von Netznutzungsentgelten für elektrische Energie und über Prinzipien der Netznutzung, Essen 2002, 23. April 2002.

Burgat, P., and C. Jeanrenaud (1990). Mesure de l'Efficacité Productive et de l'Efficacité-Cout: Cas des Dechets Menagers en Suisse. Working Paper 9002, Institut de Recherches Economiques et Regionales. Neuchatel (Valais): Université de Neuchatel.

Burns, P.J, and T. Weyman-Jones (1994a). Cost Drivers and Cost Efficiency in Electricity Distribution: a Stochastic Frontier Approach. London: Centre for the Study of Regulated Industries (CRI).

Burns, P.J., and T. Weyman-Jones (1994b). Regulatory Incentives, Privatisation, and Productivity Growth in UK Electricity Distribution. London: Centre for the Study of Regulated Industries (CRI).

Button, K., and T.G. Weyman-Jones (1992). Ownership Structure, Institutional Organisation, and Measured X-Efficiency. American Economic Review (Papers and Proceedings) 82: 439-445.

Byrnes, P. (1985). Ownership and Efficiency in the Water Supply Industry: an Application of the Non-Parametric Approach to Efficiency Measurement. Ph.D. Thesis: Southern Illinois University.

Byrnes, P., R. Fare, and S. Grosskopf (1984). Measuring Productive Efficiency: an Application to Illinois Strip Mines. Management Science 30: 671-681.

Byrnes, P., R. Fare, S. Grosskopf, and C.A.K. Lovell (1988). The Effect of Unions on Productivity: US Surface Mining of Coal. Management Science 34: 1037-1053.

Caves, D, Herriges, J and Windle, R (1990) 'Customer Demand for Service Reliability in the Electric Power Industry: A Synthesis of the Outage Cost Literature' Bulletin of Economic Research 2 (2) 79-119.

CEER (2001) Quality of Electricity Supply: Initial Benchmarking on Actual Levels, Standards and Regulatory Strategies, Council of European Energy Regulators.

CEPA (2003) 'Background to work on assessing efficiency for the 2005 distribution price control review' Cambridge Economic Policy Associates.

Chang, K.P., and P.H. Kao (1992). The Relative Efficiency of Public Versus Private Municipal Bus Firms: an Application of Data Envelopment Analysis. Journal of Productivity Analysis 3: 67-84.

Charnes, A., and W.W. Cooper (1985). Preface to Topics in Data Envelopment Analysis. Annals of Operations Research 2: 59-94.

Charnes, A., W.W. Cooper, A. Lewin, and L.M. Seiford (1994, 1995). Data Envelopment Analysis: Theory, Methodology, and Applications. Boston: Kluwer Academic Publishers.

Charnes, A., W.W. Cooper, and S. Li (1989). Using Data Envelopment Analysis to Evaluate Efficiency in the Economic Performance of Chinese Cities. Socio-Economic Planning Sciences 23: 325-344.

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, and E. Rhodes (1981). Evaluating Program and Managerial Efficiency: an Application of Data Envelopment Analysis to Program Follow Through. Management Science 27: 668-697.

Charnes, A., W.W. Cooper, J. Rousseau, and J. Semple (1987). Data Envelopment Analysis and Axiomatic Notions of Efficiency and Reference Sets. Research Report 558, Center for Cybernetic Studies, University of Texas.

Charnes, A., W.W. Cooper, L. Seiford, and J. Stutz (1982). A Multiplicative Model for Efficiency and Piecewise Cobb-Douglas Envelopments. Operations Research Letters 2: 101-103.

Christensen, L.R., and W.H. Greene (1976). Economies of Scale in U.S. Electric Power Generation. Journal of Political Economy 84: 655-676.

Christensen, L.R., D.W. Jorgenson, and L.J. Lau (1971). Conjugate Duality and the Transcendental Logarithmic Function. Econometrica 39: 255-256.

Christensen, L.R., D.W. Jorgenson, and L.J. Lau (1973). Transcendental Logarithmic Production Frontiers. Review of Economics and Statistics 55: 28-45.

Coelli, T.J. (1996a). A Guide to Frontier Version 4.1: a Computer Program for Stochastic Frontier Production and Cost Function Estimation. Discussion Paper 96/07, Department of Econometrics, University of New England, Armidale, Australia.

Coelli, T.J. (1996b). A Guide to DEAP Version 2.1: a Data Envelopment Analysis Computer Program. Discussion Paper 96/08, Department of Econometrics, University of New England, Armidale, Australia.

Coelli, T.J. (1997). A Multi-Stage Methodology for the Solution of Orientated DEA Models. Centre for Efficiency and Productivity Analysis, University of New England, mimeo.

Coelli, T.J., and S. Perelman (1996). A Comparison of Parametric and Non-Parametric Distance Functions with an Application to European Railways. CREPP Discussion Paper, University of Liège.

Cubbin, J., S. Domberger, and S. Meadowcroft (1987). Competitive Tendering and Refuse Collection: Identifying the Sources of Efficiency Gains. Fiscal Studies 8: 49-58.

Danilin, S., I. Materov, S. Rosefielde, and C.A.K. Lovell (1985). Measuring Enterprise Efficiency in the Soviet Union: a Stochastic Frontier Analysis. Economica 52: 225-234.

Dantzig, G.B. (1963). Linear Programming and Extensions. Princeton: Princeton University Press.

De Alessi, L. (1980). The Economics of Property Rights: a Review of the Evidence. Research in Law and Economics 2: 1-47.

Debreu, G. (1951). The Coefficient of Resource Utilisation. Econometrica 19: 225-234.

Deprins, D. (1983). Estimation de Frontière et Mesure d'Efficacité: Approches Descriptive et Econometrique. Revue Belge de Statistique, d'Informatique et de Recherche Operationelle 23: 19-62.

Deprins, D., L. Simar, and H. Tulkens (1984). Measuring Labor-Efficiency in Post Offices. In M. Marchand, P. Pestieau, and H. Tulkens (eds.): The Performance of Public Enterprises: Concepts and Measurement. Amsterdam: North-Holland.

Dialynas, E N, Megaloconomos, S M and Dali, V C (2001) '"Interruption cost analysis for the electrical power customers in Greece' Proceedings of the 16th CIRED conference 2001.

Doane, M J, Hartman, R S and Woo, C K (1988) 'Household preference for interruptible rate options and revealed value of service reliability' Energy Journal 9 121-34.

DTe (1999): Price Cap Regulation in the Electricity Sector, Information and Consultation Document.

DTe (2000): Guidelines for Price Cap Regulation in the Dutch Electricity Sector in the period from 2000 to 2003.

DTe (2002) ‘Maatstafconcurrentie, Regionale Netbedrijven Elektriciteit’, DTe, Den Haag, available at www.dte.nl.

DTe (2002). An overview of the first regulatory review of the regional electricity network businesses. Den Haag: DTe. Also available at www.dte.nl

Dugan, R C, McGranaghan M F, Beaty, H W (1996) Electrical Power Systems Quality, McGraw-Hill, New York.

Fare, R., S. Grosskopf, and J. Logan (1983). The Relative Efficiency of Illinois Electric Utilities. Resources and Energy 5: 349-367.

Fare, R., S. Grosskopf, and C.A.K. Lovell (1985). The Measurement of the Efficiency of Production. Boston (MA): Kluwer Academic Publishers.

Fare, R., S. Grosskopf, and C.A.K. Lovell (1994). Production Frontiers. Cambridge: Cambridge University Press.

Fare, R., S. Grosskopf, and W. Weber (1989). Measuring School District Performance. Public Finance Quarterly 17: 409-428.

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

Farrell, M.J. (1957). The Measurement of Productive Efficiency. Journal of the Royal Statistical Society, Series A, 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.

Forsund, F.R., and S.A.C. Kittelsen (1992). Efficiency Analysis of Norwegian District Courts. Journal of Productivity Analysis 3: 277-306.

Forsund, F.R., and S.A.C. Kittelsen (1997). Productivity Development of Norwegian Electricity Distribution Utilities. Resource and Energy Economics, forthcoming.

Fraser, R (1994) 'Price, Quality and Regulation: An Analysis of Price Capping and the Reliability of Electricity Supply' Energy Economics 16 (3) 175-83.

Fried, H.O., C.A.K. Lovell, and S.S. Schmidt (eds.) (1993). The Measurement of Productive Efficiency. New York: Oxford University Press.

Frontier Economics (2000). The efficiency of the Dutch network and supply companies. London: Frontier Economics. Also available at www.dte.nl

Gabszewicz, J J and Wauthy, X Y (2002) 'Quality underprovision by a monopolist when quality is not costly' Economic Letters 77 65-72.

Galal, A., et al./World Bank (1994). Welfare Consequences of Selling Public Enterprises: an Empirical Analysis. New York: Oxford University Press for the World Bank.

Gal-Or, E (1983) 'Quality and quantity competition' Bell Journal of Economics 14 590-600.

Gastaldo, E M, Capeletto, G J and Bittencourt, H R (2001) 'Statistical correlations between indicators of power supply and consumers' satisfaction' Proceedings of the 16th CIRED conference 2001.

Gates, J, Billinton, R and Wacker, G (1999) 'Electric service reliability worth evaluation for government, institutions and office buildings' IEEE Transactions on Power Systems 14 43-48.

Grabowski, R., and C. Pasurka (1988). The Relative Technical Efficiency of Northern and Southern US Farms in 1860. Southern Economic Journal 54: 598-614.

Green, A., C. Harris, and D. Mayes (1991). Estimation of Technical Inefficiency in the Manufacturing Industry. Applied Economics 23: 1637-1647.

Greene, W.H. (1993b). The Econometric Approach to Efficiency Analysis. In Fried, H.O., C.A.K. Lovell, and S.S. Schmidt (eds.) (1993). The Measurement of Productive Efficiency. New York: Oxford University Press.

Grosskopf, S., and V. Valdmanis (1987). Measuring Hospital Performance: a Non-Parametric Approach. Journal of Health Economics 6: 89-107.

Guilkey, D.K., C.A.K. Lovell, and R.C. Sickles (1983). A Comparison of the Performance of Three Flexible Functional Forms. International Economic Review 24: 591-616.

Hammond, C.J. (1992). Privatisation and Efficiency of Decentralised Electricity Generation: Some Evidence from Inter-War Britain. Economic Journal 102: 538-553.

Hausman, W.J., and J.L. Neufeld (1991). Property Rights versus Public Spirit: Ownership and Efficiency of US Electric Utilities prior to Rate-of-Return Regulation. Review of Economics and Statistics 73: 414-423.

Hebden, J. (1983). Applications of Econometrics. London: Philip Allan.

Hjalmarsson, L., and A. Veiderpass (1992a). Efficiency and Ownership in Swedish Electricity Retail Distribution. Journal of Productivity Analysis 3: 7-23.

Hjalmarsson, L., and A. Veiderpass (1992b). Productivity in Swedish Electricity Retail Distribution. Scandinavian Journal of Economics 94/Supplement: S193-S205.

Huettner, D.A., and J.H. Landon (1977). Electric Utilities: Scale Economies and Diseconomies. Southern Economic Journal 44: 883-912.

IEEE (1999) IEEE trial-use guide for electric power distribution reliability indices, IEEE Std 1366-1998.

Jamar, M.A., and H. Tulkens (1990). Mesure de l'Efficacité de l'Activité des Tribunaux et Evaluation de l'Arriére Judiciaire. CORE Working Paper 90/1, UCL, Louvain-La-Neuve.

Jamasb, T. (Department of Applied Economics, University of Cambridge), Pollitt; M. (Judge Institute of Management, University of Cambridge), 2000: „Benchmarking and Regulation of Electricity Transmission and Distribution Utilities: Lessons Learned from International Experience".

Jamasb, T and Pollitt, M (2002) 'Benchmarking and Regulation: International Electricity Experience' Utilities Policy 9 (3) 107-30.

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 19: 233-238.

Kahn, A (2002) 'The adequacy of prospective generation investments under price control mechanisms' The Electricity Journal 15 37-46.

Kalirajan, K.P., and C. Yong (1993). Can Chinese State Enterprises Perform Like Market Entities: Productive Efficiency in the Chinese Iron and Steel Industry. Applied Economics 25: 1071-1080.

Kidokoro, Y (2002) 'The Effects of Regulatory Reform on Quality' Journal of Japanese and International Economies 16 135-46.

Kim, J and Kim, J (1996) 'Quality choice of multi-product monopolist and spill-over effects' Economic Letters 52 345-52.

Kittelsen, S.A.C. (1993). Stepwise DEA: Choosing Variables for Measuring Technical Efficiency in Norwegian Electricity Distribution. Department of Economics Working Paper, University of Oslo.

Kittelsen, S.A.C. (1996). Regulating Electricity Distribution Utilities with Yardstick Competition Based on DEA. SNF Working Paper. Department of Economics, University of Oslo.

Koopmans, T.C. (1951). An Analysis of Production as an Efficient Combination of Activities. In T.C. Koopmans (ed.): Activity Analysis of Production and Allocation. Cowles Commission for Research in Economics, Monograph 13. New York: Wiley.

Kopp, R.J., and V.K. Smith (1980). Frontier Production Function Estimates for Steam Electric Generation: a Comparative Analysis. Southern Economic Journal 46: 1049-1059.

Korhonen, P and Syrjänen, M (2002) 'Evaluation of cost efficiency in Finnish Electricity Distribution' Helsinki School of Economics.

Kumbhakar, S.C. (1988). On the Estimation of Technical and Allocative Inefficiency Using Stochastic Frontier Functions: the Case of US Class I Railroads. International Economic Review 29: 727-743.

Kumbhakar, S.C., S. Ghosh, and J.T. McGuckin (1991). A Generalised Production Frontier Approach for Estimating the Determinants of Inefficiency in US Dairy Farms. Journal of Business and Economic Statistics 9 (3).

Land, K.C., C.A.K. Lovell, and S. Thore (1988). Chance-Constrained Efficiency Analysis. Working Paper, Department of Economics, University of North Carolina.

Laffont, J J and Tirole, J (1993) A Theory of Incentives in Regulation and Procurement, MIT Press, Cambridge (MA).

Langset, T, Trengereid, F, Samdal, K, and Heggset, J (2001) 'Quality Dependent Revenue Caps - A Model for Quality of Supply Regulation' Proceedings of CIRED 2001, Amsterdam.

Lehmann, E.L. (1975). Nonparametrics: Statistical Methods Based on Ranks. San Francisco: Holden Day.

Lehtonen, M and Lemstrom, B (1995) 'Comparison of the methods for assessing the customers' outage costs' Proceedings of EMPD '95.

Leibenstein, H. (1966). Allocative Efficiency vs. X-Efficiency. American Economic Review 56: 392-413.

Leibenstein, H., and S. Maital (1992). Empirical Estimation and Partitioning of X-Inefficiency: A Data-Envelopment Approach. American Economic Review (Papers and Proceedings) 82: 428-433.

Lo Schiavo, L and Malaman, R (2003) 'Regulation and Incentives for Improving Continuity of Supply: The Experience of Italy and a Comparison with other EU Countries' Proceedings of CIGRE/IEEE Symposium on Quality and Security of Supply, Montreal 8-10 October 2003.

Lovell, C.A.K. (1994). Linear Programming Approaches to Measurement and Analysis of Productive Efficiency. Top 2: 175-248.

Lovell, C.A.K. (1993). Production Frontiers and Productive Efficiency. In Fried, H.O., C.A.K. Lovell, and S.S. Schmidt (eds.) (1993). The Measurement of Productive Efficiency. New York: Oxford University Press.

Lovell, C.A.K., L.C. Walters, and L.L. Wood (1990). Stratified Models of Education Production Using DEA and Regression Analysis. Working Paper 90/5, Department of Economics, University of North Carolina.

McElroy, M. (1987). Additive General Error Models for Production, Cost, and Derived Demands or Share Systems. Journal of Political Economy 95: 737-757.

Meeusen, W., and J. van den Broeck (1977). Efficiency Estimation from a Cobb-Douglas Production Function with Composed Error. International Economic Review 8: 435-444.

Munasinghe, M (1981) 'Optimal Electricity Supply. Reliability, pricing and system planning' Energy Economics 3 140-52.

Mussa, M and Rosen, S (1978) 'Monopoly and product quality' Journal of Economic Theory 18 301-17.

Nerlove, M. (1963). Returns to Scale in Electricity Supply. In A. Zellner (ed.) (1968). Readings in Economic Statistics and Econometrics. Boston: Little, Brown and Co., 409-439.

Neuberg, L.G. (1977). Two Issues in the Municipal Ownership of Electric Power Distribution Systems. Bell Journal of Economics 8: 303-323.

Newbery, D., and M.G. Pollitt (1997). The Restructuring and Privatisation of the CEGB - Was It Worth It? Department of Applied Economics Working Paper (Published in: Journal of Industrial Economics 45/3: 269-303). Faculty of Economics, University of Cambridge.

Norman, M., and B. Stoker (1991). Data Envelopment Analysis: the Assessment of Performance. Chichester: John Wiley and Sons.

NVE: Agrell / Bogetoft (2003), Benchmarking für Regulierung, Abschlussbericht für die NVE.

Office of Gas and Electricity Markets (OFGEM), (2001) 'The Information and Incentives Project. Developing the incentive scheme. Final proposals.' London, England,

Office of Gas and Electricity Markets (OFGEM), "Developing Network Monopoly Price Controls, Update Document", February 2003, 05/03.

Office of the Regulator General, Victoria (1998): 2001 Electricity Distribution Price Review, Efficiency Measurement and Benefit Sharing. Consultation Paper No 3. December.

Pandey, M and Billinton, R (1999) 'Reliability worth assessment in a developing country – residential survey results' IEEE Transactions on Power Systems 14 1226-31.

Pestieau, P., and H. Tulkens (1990). Assessing the Performance of Public Sector Activities: Some Recent Evidence from the Productive Efficiency Viewpoint. CORE, Discussion Paper 90/60, UCL, Louvain-La-Neuve.

Pitt, M., and L.F. Lee (1981). The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry. Journal of Development Economics 9: 43-64.

Petrov Konstantin, Viren Adjodhia, Paul Giesbertz, Erland Eggen, Oddbjørn Fredriksen Micro and Macro Methods for Benchmarking in the Networks– Conference Distributech, November, 2001, Berlin.

Pollitt, M.G. (1993). The Relative Performance of Publicly-Owned and Privately-Owned Electric Utilities: International Evidence. D.Phil. Thesis, University of Oxford.

Pollitt, M.G. (1995). Ownership and Performance in Electric Utilities. Oxford: Oxford University Press.

Pollitt, M.G. (1996). Ownership and Efficiency in Nuclear Power Production. Oxford Economic Papers 48: 342-360.

Price, C., and T.G. Weyman-Jones (1993). Malmquist Indices of Productivity Change in the UK Gas Industry Before and After Privatisation. Loughborough University, Economics Research Paper 93/12, then published in Applied Economics, January 1996.

Sanghvi, A (1982) 'Economic Costs of Electricity Supply Interruptions. US and Foreign Experience' Energy Economics 4 180–98.

Sappington, D (2003) 'The Impact of Incentive Regulation on Retail Telephone Service Quality in the United States' Review of Network Economics 2 (4) 355-75.

Sappington, D and Weisman, D (1996) 'Revenue Sharing in Incentive Regulation Plans' Information Economics and Policy 8 229-48.

Scarsi, G.C. (1999). Local Electricity Distribution in Italy: Comparative Efficiency Analysis and Methodological Cross-Checking, Milan (Italy): Fondazione ENI Enrico Mattei (FEEM), Working Paper no. 16/1999.

Schmidt, P. (1984). An Error Structure for Systems of Translog Cost and Share Equations. Econometrics Workshop Paper 8309. Department of Economics, Michigan State University.

Schmidt, P. (1986). Frontier Production Functions. Econometric Reviews 4: 289-328.

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.

Seiford, L.M. (1996). Data Envelopment Analysis: The Evolution of the State of the Art. Journal of Productivity Analysis 7: 99-138.

Seiford, L.M., and R.A. Thrall (1990). Recent Developments in DEA - the Mathematical Programming Approach to Frontier Analysis. Journal of Econometrics 46: 7-38.

Sheshinski E (1976) 'Price, Quality and Quantity Regulation in Monopoly Situations' Economica 43 127-37.

Spence, A M (1975) 'Monopoly, Quality and Regulation' Bell Journal of Economics 6 417-29.

Sullivan, M and Sheehan, M (2000) 'Observed Changes in Residential and Commercial Customer Interruption Costs in The Pacific Northwest between 1989 and 1999' IEEE PES Summer Meeting 4 2290–93.

Swan, P (1970) 'The durability of consumer goods' American Economic Review 60 889-94.

Ter-Martirosyan, A (2003) 'The Effects of Incentive Regulation on Quality of Service in Electricity Markets' Dept. Of Economics, George Washington University, Working Paper, March 2003.

Thanassoulis, E., R.G. Dyson, and M.J. Foster (1987). Relative Efficiency Assessments Using Data Envelopment Analysis: an Application to Data on Rates Departments. Journal of the Operational Research Society 38: 397-411.

Thiry, B., and H. Tulkens (1992). Allowing for Technical Inefficiency in Parametric Estimation of Production Functions for Urban Transit Firms. Journal of Productivity Analysis 3: 45-65.

Train, K E (1991) Optimal Regulation, MIT Press, Cambridge (MA) and London (UK).

Tulkens, H. (1986). La Performance Productive d'un Service Public. Definitions, Méthodes de Mesure et Application à la Régie des Postes en Belgique. L'Actualité Economique, Revue d'Analyse Economique 62: 306-335.

Tulkens, H., and P. Vanden Eeckaut (1990). Productive Efficiency Measurement in Retail Banking in Belgium. CORE Working Paper, UCL, Louvain-La-Neuve.

Vanden Eeckaut, P., H. Tulkens, and M.A. Jamar (1992). Cost Efficiency in Belgian Municipalities. In Fried, H.O., C.A.K. Lovell, and S.S. Schmidt (eds.) (1993). The Measurement of Productive Efficiency. New York: Oxford University Press.

Vickers, J and Yarrow, G (1988) Privatisation: An Economic Analysis, MIT Press, Cambridge (MA).

Vickers, J.S., and G.K. Yarrow (1988). Privatization: an Economic Analysis. Cambridge (MA): The MIT Press.

Weisman, D (2002) 'Price Regulation and Quality' mimeo, Kansas State University, July.

Weyman-Jones, T.G. (1992). Problems of Yardstick Regulation in Electricity Distribution. In M. Bishop, J. Kay, and C. Mayer (1995): The Regulatory Challenge. Oxford: Oxford University Press.

White, L (1977) 'Market structure and product varieties' American Economic Review 67 179-82.

Woo, C K and Pupps R L (1992) 'Cost of service disruptions to electricity customers' Energy 17 109-26.

Zellner, A. (1968). Readings in Economic Statistics and Econometrics. Boston (MA): Little, Brown and Company.

KEMA

Annex: Regression Results

Dependent variable: LTOTEX
Current sample: 1 to 44
Number of observations: 44

Mean of dep. var.	11.3928	LM het. test	1.15434 [.283]			Estimated	Standard		
Std. dev. of dep. var.	1.39387	Durbin-Watson	1.90478 [<.481]	EQ 1	Variable	Coefficient	Error	t-statistic	P-value
Sum of squared residuals	2.42217	Jarque-Bera test	6.28455 [.043]		C	-2.56873	0.404851	-6.34487	[.000]
Variance of residuals	0.059077	Ramsey's RESET2	1.99601 [.165]		LKWH	0.053669	0.095597	0.561416	[.578]
Std. error of regression	0.243058	F (zero slopes)	686.565 [.000]		LCUST	0.98883	0.104787	9.43656	[.000]
R-squared	0.971007	Schwarz B.I.C.	-2.64151						
Adjusted R-squared	0.969593	Log likelihood	1.35626						

Dependent variable: LTOTEX
Current sample: 1 to 44
Number of observations: 44

Mean of dep. var.	11.3928	LM het. test	1.43626 [.231]			Estimated	Standard		
Std. dev. of dep. var.	1.39387	Durbin-Watson	1.78305 [<.387]	EQ 2	Variable	Coefficient	Error	t-statistic	P-value
Sum of squared residuals	2.27497	Jarque-Bera test	5.50491 [.064]						
Variance of residuals	0.056874	Ramsey's RESET2	1.2032 [.279]		C	-2.65543	0.400869	-6.62417	[.000]
Std. error of regression	0.238483	F (zero slopes)	476.302 [.000]		LKWH	0.146591	0.110154	1.33078	[.191]
R-squared	0.972769	Schwarz B.I.C.	-2.61821		LCUST	0.789232	0.161132	4.89804	[.000]
Adjusted R-squared	0.970727	Log likelihood	2.7356		LCML	0.073967	0.045977	1.60878	[.116]

Dependent variable: LTOTEX
Current sample: 1 to 44
Number of observations: 44

Mean of dep. var.	11.3928	LM het. test	0.78744 [.375]			Estimated	Standard		
Std. dev. of dep. var.	1.39387	Durbin-Watson	2.01961 [<.630]	EQ 3	Variable	Coefficient	Error	t-statistic	P-value
Sum of squared residuals	2.37569	Jarque-Bera test	8.09472 [.017]		C	-2.45189	0.374014	-6.55562	[.000]
Variance of residuals	0.057944	Ramsey's RESET2	1.67088 [.204]		LCUST	0.988007	0.061005	16.1955	[.000]
Std. error of regression	0.240715	F (zero slopes)	700.398 [.000]		LCML	0.041885	0.039516	1.05995	[.295]
R-squared	0.971563	Schwarz B.I.C.	-2.66089						
Adjusted R-squared	0.970176	Log likelihood	1.78252						