Estimating the cost of improving quality in electricity distribution: A parametric distance function approach

Tim Coelli (*University of Queensland*), Axel Gautier (*University of Liège*)
Sergio Perelman (*University of Liège*) and Roxana Saplacan-Pop (*Electricité de France*)

Abstract

In this paper we are interested in the way electricity distribution operators anticipate and prevent potential outages by increasing maintenance activities and/or capital investments. We make use of the parametric distance function approach, assuming that outages enter in the firm production set as an input, an imperfect substitute for maintenance activities and investment. This allows us to identify the sources of technical inefficiency and the underlying trade-off faced by operators between quality and other inputs and costs. For this purpose, we use panel data on 92 electricity distribution units operated by ERDF (*Electricité de France - Réseau Distribution*) in the 2003–2005 financial years. Assuming a multi-output multi-input translog technology, it appears that technical efficiency is positively correlated with the share of underground lines and with the age of capital. Moreover, the results show that shadow price of quality varies dramatically: from 2.7 € to 15.7 €, by customer interrupted among the operators. Furthermore, as one would expect, marginal quality improvements tend to be more expensive as a network approaches 100% reliability.

Corresponding author:
Sergio Perelman
CREPP, HEC Management School
University of Liège
Bd. du Rectorat 7 (B31)
4000 Liège, Belgium
Phone: +32 43663098
Fax: +32 43669318
Email: sergio.perelman@ulg.ac.be

Introduction

The frequency and the duration of power outages are the two key measures of quality that electricity distribution utilities pay particular attention to. Other than direct costs of outages, represented by opportunity costs and repair expenditures, firms operating in a regulated framework also risk penalties, generally a fixed amount for each customer affected by long duration outages (CEER, 2008). To prevent outages and these related costs, operators have two main possibilities, either to increase maintenance or to make new investments, e.g. replace overhead lines by underground lines. In this paper we are mainly interested in this issue and, more precisely, in the way that electricity distribution operators anticipate and prevent potential outages by increasing maintenance activities and/or capital investments.

We make use of the parametric distance function approach proposed in the activity analysis literature to deal with undesirable outputs (Färe et al., 1993). The same approach is applied here, but instead of assuming that outages are an undesirable output, we assume that they enter in the firm production set as an input, i.e., that outages are an imperfect substitute for maintenance activities and investment. Therefore, following Growitsch et al. (2005), we postulate that the corresponding distance function is input oriented. This allows us to identify the underlying trade-off faced by operators, between quality and other inputs and costs.

In this study we use panel data on 92 electricity distribution units operated by ERDF (*Electricité de France - Réseau Distribution*) in France in the 2003–2005 financial years. Compared with similar studies, we have access to very comprehensive and comparable data, mainly on the value of capital. This database allows us to estimate a flexible translog multi-output multi-input technology. On the output side, we chose a specification that takes into account the main output dimensions of the electricity distribution activity: i) the number of customers; ii) the surface area served and; iii) the GWh of electricity distributed. On the input side, the three dimensions retained are: i) operational expenditures; ii) capital; and iii) quality, represented by the number of interruptions (longer than 3 minutes in duration).

Given the flexible nature of the translog distance function, we use for computation purposes a stochastic frontier approach (SFA) and a parametric (deterministic) linear programming approach (PLP). Both approaches give similar results, on average. For further analysis we select the parameters and the results obtained from PLP, as this approach allows us to impose restrictions implied by economic theory, in a very simple way, on the parameters of the distance function, such as monotonicity.

On the one hand, our results indicate that SFA technical efficiency (TE) is positively correlated with the share of underground lines and with the age of capital. On the other hand, using the computed PLP translog parameters, several measurements are done that allow us to describe the main characteristics of the underlying production technology. Among others, the distance function elasticities with respect to inputs and outputs at each point of the boundary surface. And using these measurements, shadow prices can be derived, for the quality (outages) measures.

These results are potentially useful for the operators themselves, who can obtain information regarding the marginal cost of reducing interruptions. They are also useful for regulators, who could use them for the design of incentive schemes that incorporate quality measures. Moreover, the results show that shadow price of quality varies dramatically: from 2.7 € to 15.7 €, by customer interrupted among the operators. Furthermore, as one would expect, marginal quality improvements tend to be more expensive as a network approaches 100% reliability.

The remainder of the paper is organized as follows. In Section 2 we survey the literature on benchmarking analyses in electricity distribution including service quality while Section 3 describes the electricity distribution sector in France. Sections 4 and 5 present the methodology and the data used in estimation, respectively. In Section 6 we report the main results of this study and in Section 7 we draw some conclusions.

2 Literature

Most benchmarking analyses in electricity distribution have involved models that incorporate standard output characteristics, such as energy supplied (in GWh), number of customers and network size (e.g., service area or network length). For example, see the literature review in London Economics (1999) and Jamasb and Pollitt (2001). Very few studies have included quality of service measures in these models. Exceptions are the studies by Giannakis, Jamasb and Pollitt (2005), Growitsch, Jamasb and Pollitt (2005), Coelli et al (2007) and Jamasb, Orea and Pollitt (2010).

Giannakis et al (2005) use data envelopment analysis (DEA) methods to measure technical efficiency (TE) and total factor productivity growth (TFP) in 14 UK distribution authorities over the 1991/92 to 1998/99 period. The DEA method is used to estimate a non-parametric input distance function that involves three output variables (energy supplied, customers and network length). Four models involving different input sets are considered: (i) operating expenditure (OPEX); (ii) total expenditure (TOTEX); (iii) number of interruptions (NINT) and total interruptions (TINT); and (iv) TOTEX, NINT and TINT. They find that the TE scores of the various models are positively (but not perfectly) correlated, and that the TE scores rise when the NINT and TINT quality variables are added to the TOTEX model (a result that is to be mathematically expected when variables are added to a DEA model).[1]

Growitsch et al (2005) use stochastic frontier analysis (SFA) methods to estimate an input distance function using data on 505 electricity distribution utilities from eight European countries in the 2002 financial year. Their models contain two output variables (energy supplied and customers) and either one input variable (TOTEX) or two input variables (TOTEX and TINT). They use the Battese and Coelli (1995) SFA model to investigate the effects of customer density (customers per network km) and country (using dummy variables) upon technical efficiency scores. They find that the inclusion of the quality variable reduces TE for all but the large firms, plus they find that the TE scores from the two models are significantly negatively correlated, both findings being in contrast to those of Giannakis et al (2005).

Jamasb et al (2010) estimate the marginal cost of quality improvements of 12 UK distribution companies for the period 1995-2003. For that, they run fixed-effect estimations of the link between the cost of electricity distribution (identified with TOTEX or CAPEX) and a series of cost drivers including the energy delivered, the network length, the network energy losses, the customers minute lost and a time trend. They found that the cost of reducing energy losses is positive and, in average, equal to 2.8 pence per kWh. This marginal cost of improving quality is smaller than the penalty/reward set by the regulator (4.8 pence per KWh) for lower/higher delivered quality. Moreover, the marginal cost of improving quality increases with the quality delivered.

The above studies are to be commended for introducing quality variables into these benchmarking models. However, these studies contain some shortcomings. First, they all make use of TOTEX measures which contain capital expenditure (CAPEX) measures which need not reflect the actual amount of capital services consumed in a particular year. Second, the UK studies suffer from small sample size problems while the inter-country study suffers from difficulties associated with deflating monetary values of TOTEX in order to obtain comparable measures of implicit input usage in each country.

In the current study we aim to address these problems by making use of a detailed database on the activities of electricity distribution units operated by ERDF Réseau Distribution in France in the 2003–2005 financial years. With these data we thus avoid the small sample size problem; we avoid the international comparability problem; and we also have access to comprehensive and comparable data on the value of capital items, so we can avoid the need to use CAPEX to measure capital input services. Coelli et al (2007) used a previous version of the same data and similar methodological strategy, but relying on a comparison of parametric and non parametric approaches.

3 **Electricity distribution in France**

In France, most electricity distribution grids are owned by municipalities, individually or grouped in communities. Municipalities are in charge of the public service of electricity distribution, which they delegate to a third party, the distribution system operator (DSO), within the framework of a concession. The concession contracts between parties follow a similar model. The public service requirements are, indeed, the same all over the country.

[1] This is also seen in a DEA study by Korhonen and Syrjänen (2003) of Finnish electricity distribution operators, where the inclusion of a TINT variable into the DEA model led to increases in technical efficiency for a number of firms. For example, see their Figure 3. However, note that these results need to be treated with caution because their DEA model did not include a capital measure, which could lead to substantial biases.

The concession contracts define the rights and obligations of the distributor regarding quality of supply, customers' connections and environmental conditions. These contracts state that the distributor is remunerated by the tariff applied to final users, which is supposed to cover operating costs and investments. This tariff is the same for all the concessions (one single price for all the customers in France) and for all DSOs. The rates for the use of public electricity grids, including transmission and distribution networks, are set by the French Regulator, the CRE (Commission de Régulation de l'Energie).

Most of the municipalities delegate the management of their network to ERDF, a subsidiary of EDF, the historical electricity operator which is now publicly listed company. ERDF covers more than 95% of the territory and the remaining part is covered by local public companies.

There are 92 local distribution units. These units, known as Centers, are grouped in 23 URE (*Unité Réseau Electrique*) and further aggregated in 8 regions. During the sample period, centers were autonomous (within limits) for taking decisions regarding capital and operational expenditures. In 2006, the company was reorganized and the decision power moved to the URE, with the Centers remaining as administrative units.

The quality of electricity distribution is regulated by the CRE. The quality is measured by the minutes of interruption from which a series of exceptional events are removed (criterion B). The regulator sets a quality target and rewards/penalties are set according to the fulfilment of the objectives.

4 Methodology

We model the production process using a multi-input, multi-output input distance function and introduce the quality variable as an input variable. The logic associated with including the quality variable as an input variable is that the operators can substitute between regular inputs (labour, capital etc.) and the inconvenience faced by the customers (interruptions). The rational operator will look at the "price of interruptions" (e.g., the penalty imposed by the regulator) and compare it with the price of other inputs (e.g., labour) before deciding upon the optimal (cost minimising) mix of inputs to use.

If the production technology (frontier) is known (which is rarely the case) we can measure the distance that each data point (firm) lies below the frontier by calculating the amount by which the input vector ($\mathbf{x}$) can be proportionally reduced while holding the output vector ($\mathbf{y}$) constant. That is, for each data point ($\mathbf{x},\mathbf{y}$) we seek to find the smallest possible value of the scalar θ such that ($\theta\mathbf{x},\mathbf{y}$) remains within the feasible production set bounded by the frontier. This is illustrated (for the case of a 2-input technology) in Figure 1, where the distance that firm A is inside the frontier is equal to θ=0*B*/0*A*. This distance (i.e., technical efficiency score) equals approximately 0.7 in this diagram, suggesting that the firm could reduce input usage by 30% and still produce the same output vector.

INSERT FIGURE 1

In reality, the production frontier is rarely known. Instead it is estimated using sample data on a number of firms. This generally involves fitting an empirical frontier that aims to minimise these distances so that the frontier is a "tight-fit" to the data. In this paper we use parametric methods to estimate an input distance function.

The input distance function may be defined on the input set, $L(\mathbf{y})$, as:

$$D_I(\mathbf{x},\mathbf{y}) = \max\{\rho : (\mathbf{x}/\rho) \in L(\mathbf{y})\}, \tag{1}$$

where $\rho = 1/\theta$ and the input set $L(\mathbf{y})$ represents the set of all input vectors, $\mathbf{x} \in R_+^K$, which can produce the output vector, $\mathbf{y} \in R_+^M$. That is,

$$L(\mathbf{y}) = \left\{\mathbf{x} \in R_+^K : \mathbf{x} \text{ can produce } \mathbf{y}\right\}. \tag{2}$$

$D_I(\mathbf{x},\mathbf{y})$ is non-decreasing, positively linearly homogeneous and concave in $\mathbf{x}$, and increasing in $\mathbf{y}$. The distance function will take a value which is greater than or equal to one if the input vector, $\mathbf{x}$, is an element of the feasible input set, $L(\mathbf{y})$. That is, $D_I(\mathbf{x},\mathbf{y}) \geq 1$ if $\mathbf{x}\in L(\mathbf{y})$. Furthermore, the distance function will take a value of unity if $\mathbf{x}$ is located on the inner boundary of the input set.

Stochastic frontier analysis (SFA)

Following Coelli *et al* (2003), a translog input distance function for the case of M outputs and K inputs is specified as

$$\begin{aligned}\ln D_i = \alpha_0 + \sum_{m=1}^{M} \alpha_m \ln y_{mi} + \frac{1}{2}\sum_{m=1}^{M}\sum_{n=1}^{M} \alpha_{mn} \ln y_{mi} \ln y_{ni} + \sum_{k=1}^{K} \beta_k \ln x_{ki} \\ + \frac{1}{2}\sum_{k=1}^{K}\sum_{l=1}^{K} \beta_{kl} \ln x_{ki} \ln x_{li} + \sum_{k=1}^{K}\sum_{m=1}^{M} \delta_{km} \ln x_{ki} \ln y_{mi}, \quad i = 1,2,\ldots,N\end{aligned} \qquad (3)$$

where i denotes the i-th firm in the sample of N firms.[2] Note that to obtain the frontier surface (i.e., the transformation function) one would set D_i=1, which implies the left hand side of equation (3) is equal to zero.

Imposing homogeneity of degree +1 in inputs and rearranging we obtain

$$\begin{aligned}\ln(1/x_{Ki}) = \alpha_0 + \sum_{m=1}^{M} \alpha_m \ln y_{mi} + \frac{1}{2}\sum_{m=1}^{M}\sum_{n=1}^{M} \alpha_{mn} \ln y_{mi} \ln y_{ni} + \sum_{k=1}^{K-1} \beta_k \ln x_{ki}^* \\ + \frac{1}{2}\sum_{k=1}^{K-1}\sum_{l=1}^{K-1} \beta_{kl} \ln x_{ki}^* \ln x_{li}^* + \sum_{k=1}^{K-1}\sum_{m=1}^{M} \delta_{km} \ln x_{ki}^* \ln y_{mi} - \ln D_i, \quad i = 1,2,\ldots,N\end{aligned} \qquad (4)$$

where $x_{ki}^* = x_{ki} / x_{Ki}$.

The restrictions required for homogeneity of degree +1 in inputs are

$$\sum_{k=1}^{K} \beta_k = 1$$

and

$$\sum_{l=1}^{K} \beta_{kl} = 0, \quad k=1,2,\ldots,K, \quad \text{and} \quad \sum_{k=1}^{K} \beta_{km} = 0, \quad m=1,2,\ldots,M, \qquad (5)$$

and those required for symmetry are

$$\alpha_{mn} = \alpha_{nm}, \quad m,n=1,2,\ldots,M, \quad \text{and} \quad \beta_{kl} = \beta_{kl}, \quad k,l=1,2,\ldots,K. \qquad (6)$$

To estimate this model using SFA methods we replace the distance term with an error term that has two *i.i.d.* components. That is, we set $-\ln D_i = v_i - u_i$, where $v_i \sim |N(0,\sigma_v^2)|$ is a symmetric error to account for data noise and the u_i is a one-sided error to account for technical inefficiency. The technical efficiency score for the i-th firm is predicted using the conditional expectation: $E[\exp(-u_i \mid v_i - u_i)]$, which takes a value between 0 and 1. The model is estimated using maximum likelihood (ML) methods. Note that prior to estimation the variance parameters, σ_v^2 and σ_u^2 are re-parameterised as $\gamma = \sigma_u^2/(\sigma_u^2 + \sigma_v^2)$ and $\sigma^2 = \sigma_u^2 + \sigma_v^2$ for computational convenience.

The u_i term is often modelled as a truncated normal distribution, of the form $u_i \sim |N(\mu_i, \sigma_u^2)|$. However, in this study we make use of the more generalised model proposed by Battese and Coelli (1995), which allows one to investigate the effects of various factors upon efficiency levels. In this model the inefficiency term is made an explicit function of a vector of exogeneous characteristics, $\mathbf{z}_i$, by specifying that the u_i are independently (but not identically) distributed as non-negative truncations of a general normal distribution

---

[2] Note that in our application we have annual data on 92 units over a three year period. Hence we have 276 observations. Given the short time period we assume that there has been no technological progress over this period and hence pool the data as if it was a single year of data on 276 firms when estimating the production frontiers.

$$u_i \sim |N(m_i, \sigma_u^2)|, \tag{7}$$

where $m_i = \delta_0 + \sum_{j=1}^{J} \delta_j z_{ij}$ and the δ_0 and δ_j are unknown parameters to be estimated.

Within this framework, the values of the unknown parameters in (4) and (7) are obtained simultaneously using maximum likelihood estimation. The expressions for the likelihood function and first partial derivatives are presented in Battese and Coelli (1993), as well as the expression for $E[\exp(-u_i \mid v_i - u_i)]$.

Shadow prices

Furthermore, as well as measuring the effect of quality upon TE scores, we also make use of the methods described in Grosskopf et al (1995) and Coelli and Rao (1998) to derive measures of the shadow price of quality from the curvature of the estimated distance functions. This information could be quite valuable in allowing one to assess the degree to which rewards for quality outcomes could influence the services provided.

Shadow price information is obtained using the method outlined in Grosskopf et al (2005), Morrison Paul and Nehring (2005), and others. That is, we obtain ratios of shadow prices from the ratios of derivatives of the input distance function as:

$$\frac{v_{ik}}{v_{il}} = \frac{\partial D_i / \partial x_{ik}}{\partial D_i / \partial x_{il}} \tag{8}$$

for the case of inputs, and

$$\frac{u_{im}}{u_{in}} = \frac{\partial D_i / \partial y_{im}}{\partial D_i / \partial y_{in}} \tag{9}$$

for the case of outputs.

In order to compute ratios of shadow prices, we compute input distance partial elasticities with respect to inputs:

$$s_{ki} = \frac{\partial \ln D_i}{\partial \ln x_{ki}} = \beta_k + \sum_{l=1}^{K} \beta_{kl} \ln x_{li} + \sum_{m=1}^{M} \delta_{km} \ln y_{mi} \ , \tag{10}$$

and with respect to outputs:

$$r_{mi} = \frac{\partial \ln D_i}{\partial \ln y_{mi}} = \alpha_m + \sum_{n=1}^{M} \alpha_{mn} \ln y_{ni} + \sum_{k=1}^{K} \delta_{km} \ln x_{ki} \ . \tag{11}$$

These elasticities have also a direct interpretation as shadow shares. Shares of inputs in total input, for s_{ki}, and shares of outputs in total output, for r_{mi}. Moreover, combining r_{mi} we compute scale elasticities at each point:

$$e_i = \sum_{m=1}^{M} r_{mi} \ , \tag{13}$$

with $e_i > -1$, $e_i = -1$ and $e_i < -1$ indicating decreasing, constant or increasing returns to scale, respectively.

Furthermore, a well-behaved production function must satisfy some desirable properties, among them monotonicity and curvature conditions. Monotonicity implies that the input distance function analyzed here has to be non-decreasing in inputs ($s_{ki} \geq 0$) and non-increasing in outputs ($r_{mi} \leq 0$) (Färe and Primont, 1995). Curvature conditions imply that the input distance function satisfy convexity in outputs and quasi-convexity in inputs.

Unfortunately, we are unable to impose these well-behaved conditions on the SFA estimation using traditional econometrics techniques. The main reason is that these conditions cannot be introduced as simple restrictions on parameters. As a consequence, potentially the estimations will show monotonicity and curvature violations at specific points. In other words, incorrect computed values for shadow shares and shadow prices ratios at particular points.[3]

Parametric linear programming (PLP)

In this paper, we proceed in two steps. After verifying that SFA results satisfy monotonicity restrictions for average values but not for extreme points, we recomputed the input distance function using a parametric linear programming approach (PLP).

Values of unknown parameters in equation (3) are obtained by using LP as follows:

$$\text{Min} \sum_{i=1}^{N} \ln D_i \ ,$$

subject to the constraints that:

$$\ln D_i \geq 0, \quad i = 1, 2, \ldots, N,$$

$$s_{ki} \geq 0, \quad i = 1, 2, \ldots, N, \ k = 1, 2, \ldots, K,$$

$$r_{mi} \leq 0, \quad i = 1, 2, \ldots, N, \ m = 1, 2, \ldots, M,$$

as well as to the same homogeneity and symmetry constraints in (4) and (5).

5 **Data**

The selection and measurement of input and output variables is a key aspect of any efficiency analysis study. In this study we have drawn upon our knowledge of the key cost drivers in the French electricity distribution industry, along with reviewing the experiences gained in previous analyses. For example, see those studies surveyed in London Economics (1999) and Jamasb and Pollitt (2001), and more recent studies, such as Lawrence and Diewert (2006) and Edvardsen et al (2006).

Three output variables are used in the present study: energy supplied, number of customers and the service area. The amount of energy supplied in giga-watt hours (GWH) is generally the first output variable thought of, since the aim of a distribution company is to “supply electricity to customers”. Although a distribution network operator cannot normally determine the amount of electricity distributed, it has to ensure that all its network assets have the capacity to deliver this energy to its customers. Hence, the total amount of energy supplied may be viewed as a proxy for the load capacity of the network. The measure used in this study is gross electricity distributed (which includes losses).

The number of customers (CUST) is also used as an output variable in our model because we believe that this variable is needed to ensure that the model does not unfairly discriminate against those operators which sell smaller amounts of energy per customer. Furthermore, a large part of distribution activities (relating to metering services, customer connections, customer calls, etc.) are directly correlated to the number of customers. Note that our measure only includes Low Voltage (LV) customers, since industrial customers who are connected to the Medium Voltage (MV) network are rather small in number.

The surface covered in square kilometres (KM2) is a measure of network dispersion. A lot of network operations, such as routine maintenance, overhaul, vegetation management for overhead lines, etc. are closely linked to the length of MV and LV lines or, indirectly, to the size of the area served. Moreover, the reliability of a distribution network and therefore the level of quality of supply is often affected by the length of feeders, in other words, by customers’ density. In big cities, where the feeders are mostly short and underground, the number of outages should be lower than in less dense areas which tend to have a high proportion of overhead lines. As a consequence, the costs of repairs are expected to differ between urban and rural areas.

[3] See O’Donnell and Coelli (2005) for an application of a Bayesian approach to impose regularity conditions.

The net effect of using these three output variables in our model is to ensure that the key aspects of output heterogeneity are captured, so that when we conduct benchmarking comparisons using technical efficiency (TE) measures, we are conditioning on these factors and hence comparing like with like. That is, not comparing distribution units like Lille with the Southern Alps, and so on. Nevertheless, we are aware that with three output variables, we are unable to control for all environmental differences that could influence costs, such as influence of forests and mountainous terrain, ages of the assets, accessibility of lines or substations, climatic factors, etc.

The inputs used in electricity distribution are many and varied. In terms of capital inputs there are underground and overhead lines of various voltage levels, transformers, vehicles, computers, and so on. Plus we have various types of labour – technicians, engineers, managers, etc. – plus a variety of other materials and services. One could perhaps define dozens of input variables, but degrees of freedom limitations in the production model prevent us from doing that. Instead we have chosen to define only two input variables – capital inputs (CAP) and non capital inputs (OPEX).

Capital is measured using gross (not depreciated) value. We have chosen gross in preference to net because we wish to avoid the situation where an operator that has conducted a lot of recent investment is labelled as inefficient because their net capital stock is high relative to others. In using this measure we implicitly make two assumptions. First we assume that asset age does not significantly affect service potential. Second we assume that all operators have assets with similar life spans and hence that annual service potential is proportional to the stock. These assumptions are arguably quite reasonable in the current study, since all the data come from a single distribution operator (ERDF) who defines and manages very similar policies for investment, operations and network asset development across the various local distribution units.[4]

In terms of non-capital inputs, we use network operating expenses net of depreciation and interest as our aggregate measure of these items. These are the direct operational costs of local distribution units, excluding centralized network service support and overhead costs. These operational costs relate to day-to-day operations, such as:

- operating, developing and maintaining distribution network assets: looking after substations and overhead lines, fault repairs, remote control and dispatching, and so on;
- running connections services;
- providing meter services and any other customer interventions;
- relations with local authorities and customers; etc.

We could have chosen to split this OPEX grouping into labour and non-labour groups, but given that labour expense dominates this category and that outsourcing is blurring the boundaries between these two categories, we decided to use a single variable.[5]

Finally, quality is measured as the total number of interruptions (NINT) – excluding short interruptions of three minutes or less.

The total number of interruptions NINT has been calculated as follows:

$$\text{NINT} = \text{SAIFI} \times \text{Total number of customers.}$$

According to the international standards relative to quality of supply, SAIFI (System Average Interruption Frequency) is the average number of sustained interruptions (>3 min) experienced per customer served per year:

$$\text{SAIFI} = \frac{\text{Total number of customer interruptions}}{\text{Total number of customers served}}.$$

Therefore, NINT represents the total number of outages. It includes unplanned interruptions, even those for

[4] The gross capital value is computed by ERDF using replacement values, with the exception of capital materials that have reached the end of their depreciation period (expected potential service life), in which case the gross purchase value is only adjusted for inflation.

[5] CAPITAL and OPEX variables are expressed in 2005 prices using a gross industrial commodities price deflator.

which the distribution company is not responsible (e.g., due to transmission network outages), and also planned interruptions (e.g., to accommodate extensions, upgrades, etc.).

Explaining efficiency variations

We define a number of variables that can be used to investigate some of the reasons for variations in efficiency across different DSOs. The variables that we consider are as follows:

1. D2004 and D2005 are dummy variables that attempt to capture factors that vary from one year to another, such as the effects of temperature variations on demand patterns and the effects of storm events on outages.
2. UNDERG is the proportion of the network that is located underground (as opposed to being overhead on poles). We expect that the higher asset values for underground lines in CAP will be offset by the reduced maintenance requirements in OPEX and the reduced number of outages. However, there may be some other aspect to undergrounding that we have not captured in our model, and hence we include this variable to see if we can identify an additional effect.
3. DENSE is the proportion of customers that are located in towns involving less than 10,000 inhabitants.
4. AGE is the ratio of net book value to gross book value of assets. Hence it is an index of average asset age that varies between 0 and 1, with higher values indicating newer assets.
5. GROWTH is the ratio of customer numbers in the previous year to customer numbers in the current year.
6. HVCON is the amount of high voltage capacity that is contracted to industrial customers divided by total transformer potential. It is an indicator of the degree to which industrial customers are important to the DSO.
7. EXNINT is the proportion of NINT that is due to exceptional events.
8. EXMINT is the proportion of MINT (minutes of interruptions) that is due to exceptional events.

Descriptive statistics

The units of observation are the 92 ERDF Centres (Paris is not included in this study). All the values reported are in averages for the three-year period 2003-2005. Table 1 provides an overview of outputs, inputs and environmental factors. It illustrates the range of variation among Centres, not only on size, measured by the number of residential customers and the surface served, but particularly in terms of the share of underground lines and of small towns, as well as the percentage of outages, frequency and duration, due to exceptional events.

INSERT TABLE 1

Table 2 presents ratios obtained combining outputs and inputs quantities by customers' density quintiles. On the one hand, electricity consumption by customer (GWH/CUST) is on average invariant across quintiles but, as expected, capital density (CAP/KM2) varies dramatically following the evolution of customers' density (CUST/KM2). On the other hand, operational costs by customer (OPEX/CUST) diminishes from 91.4 € to 58.9 € from the first to the fifth quintiles, while the frequency of interruptions (SAIFI=NINT/CUST) varies in a similar manner, it is close to 1.5 per customer per year among Centres in the low density quintile and close to 1.0 in the highest quintile.

These observations might be seen as indicating that costs are mainly driven by the level of outages. However, the distance function estimates presented later show that it is dangerous to look at a few measures in isolation, and that the relationships are much more complex. The direction and the importance of these relations will depend, among others, on the complementarity/substitutability between OPEX, capital investments and quality.

INSERT TABLE 2

6 **Results**

In this section we report the estimates obtained using Stochastic Frontier Analysis (SFA) including the effects of the environmental variables. We compare the parameter estimates and the technical efficiency (TE) scores with those obtained computing the same distance function model using Parametric Linear Programming (PLP) with monotonicity restrictions imposed. Finally, we report partial elasticities and quality shadow prices computed for the PLP model.

Table 3 presents parameters for both the SFA and the PLP models. Note that output (y_m, $m=1, \ldots, M$) and input (x_k, $k=1, \ldots, K$) variables are in logarithms and also in deviations with respect to means and environmental factors (z_j, $j=1, \ldots, J$) in deviations with respect to means, except for dummy variables (z_1 and z_2).

INSERT TABLE 3

Note that y_1=CUST; y_2= KM2; y_3=GWH; x_1=OPEX; x_2=CAP; x_3=NINT. In SFA model, x_1 was chosen as the reference variable to impose homogeneity of degree + 1.[6] Therefore, x_1 becomes the dependent variable and x_k are replaced by $x^*_k=x_k-x_1$ for $k = 2, \ldots k$.

The results presented in Table 3 can be summarized as follows:

- Given that variables are expressed in logarithmic deviations from mean values, first order coefficients associated with outputs and inputs may be interpreted as distance function elasticities with respect to outputs and inputs at the sample mean, respectively. In both models these coefficients have the expected sign, negative for outputs elasticities (r_{mi}) and positive for inputs (s_{ki}), and are very close each other. The only exception are the coefficients associated with the number of customers (y_1=CUST) and with energy supplied (y_3=GWH) that are lower and higher, respectively, under the PLP model with respect with SFA.
- Under the SFA model, in most cases these coefficients are statistically significant, with t-ratio tests higher than 1.7. Second order terms are significant for the squared capital variable (x_2=CAP) and for the squared surface output (y_2=KM2), but insignificant for most of the other terms.
- The sum of the output elasticities provides information on returns to scale. For the SFA model we see that the sum e_i = -0.964, implying increasing returns to scale at the sample mean. That is, a 1% increase in outputs can be achieved using a 0.964% increase in inputs. For the PLP model, the results are on the other way around, e_i = -1.030, that is decreasing returns to scale at the sample mean.
- In the SFA model γ=0.174. This implies that the error term is primarily associated with noise.
- Among the environment factors (z_j variables), we note that DENSE, GROWTH, HVCON, EXNINT and EXMINT, are statistically insignificant (the 5% level). UNDERG and AGE are the only two variables that are significant. Note that a positive (negative) coefficient corresponds to a decreasing (increasing) effect on technical efficiency.
- The coefficient of UNDERG (percentage of underground lines) is negative. As expected, this factor has a negative effect on inefficiency, because underground lines are less susceptible to storm damage, etc. The marginal effect indicates that an increase of 10 percentage points in underground lines implies a decrease of near 2.5 percentage TE points in average.
- The coefficient of AGE (the ratio of net (depreciated) to gross capital in book values) is positive. It was expected that this factor would have a negative effect on inefficiency, because we expected that newer assets would require less maintenance. However, likely newer assets have more "teething problems" and hence require extra adjustments in the early years.

Summing up, the coefficients reported in Table 3 show close results between SFA and PLP models at mean sample values. That is, independent of their stochastic and deterministic nature and the fact that SFA takes simultaneously into account the potential effect of environmental variables.

Technical efficiency (TE)

Table 4 reports average technical efficiency (TE) scores for the SFA and PLP models by quintile. ERDF

[6] Results are insensitive to the choice of the reference variable, as illustrated in Coelli and Perelman (1996).

Centres are classified in quintiles by customers' density, quality (SAIFI) and underground lines (%). As expected, SFA technical efficiency scores are higher than PLP scores, 0.897 vs. 0.828 on average.[7] This is due to the role played by noise under the SFA model, as indicated before. The Pearson correlation between both scores is 0.708.

In both cases, SFA and PLP, TE scores show the same evolution, increasing parsimoniously across Centres on behave of customers' density, better quality (SAIFI) and with the percentage of underground lines. This confirms that TE in energy distribution is mainly drove by customers' density and quality considerations. Even if underground lines imply huge capital investment, compared with surface lines, at the end of the day extra capital costs are compensated by diminishing OPEX costs and frequency of interruptions. In order to identify these relationships among inputs, we turn now to a more deeper study of the underlying production technology, looking to distance function elasticities and shadow prices at all points (92 Centres, 3 years). For this purpose, we rely exclusively on PLP results. Unfortunately, but as expected, the estimated SFA technology did not verify all monotonicity restrictions at all points.

INSERT TABLE 4

Input and output distance function elasticities and shadow price ratios

Table 5 contains information on distance function elasticities with respect to inputs (s_{ki}) and outputs (r_{mi}). They are computed using equations (10) and (11) and correspond to "output shares" and to "input shares, respectively. In both cases, partial elasticities vary systematically with customer density.

On the one hand, capital share decreases while the share of operational expenditures and quality increases. As expected, partial elasticities are higher among units operating with proportionally lower quantities of a given input resource, and vice versa.

On the other hand, distance function elasticities with respect to the number of customers increases dramatically with density while, simultaneously, surface and energy distributed elasticities decrease. Summing up, scale elasticities go from increasing returns in low density units to decreasing returns in high density units.

INSERT TABLE 5

We turn now to shadow price ratios. Input shadow ratios (eq. 8) correspond to marginal rates of substitution and output shadow ratios (eq. 9) to marginal rates of transformation. Table 6 reports shadow price ratios by customer density quintiles.

On the input side the three ratios represented increase with customers' density. The first two columns illustrate the trade-off between quality and other costs, OPEX and capital. More precisely, one customer interruption (>3 minutes) has a shadow price of 4.9 € of OPEX costs or 97.1 € of gross capital investments for a low density Centre, while it cost 7.5 € and 613.5 € for a high density (urban area) Center. At the same time, the marginal rate of substitution between operational costs and capital (OPEX/CAP) increases from 20.5 € to 75.7 €. Given an average asset age ratio of 0.62 and a weighted average cost of capital (WACC) of 8% approximately, the expected OPEX/CAP ratio is close to 20. Therefore, values higher than 20 likely imply an overuse of capital in more dense areas.

In the output side, as expected, we observe a huge increase of the surface (KM2) shadow price, with respect to costumers and GWh outputs, among high customer density Centres. At the same time, the marginal rate of transformation between customers and GWh is divided by ten when the first and fifth quintiles are compared.

Moreover, in some cases shadow price ratios can be converted into shadow prices. It is possible if one can reasonably assume that the observed price of one input variable equals its shadow price. For example, if we assume this for OPEX we can therefore conclude from our PLP results that the average shadow price of quality (NINT) is 5.1 € (because the price of a unit of OPEX is one Euro). That is, it will cost approximately Five Euros to reduce the number of customer interruptions by one.

[7] The distance measures derived from the estimation of input distance functions are, by definition, equal or higher than 1.0. For presentation purposes, we transform them into technical efficiency scores, with values between 0 and 1.0, by taking the reciprocal.

INSERT TABLE 6

This shadow price information is also reported in Figure 2. The horizontal axis corresponds to quality (SAIFI) and the dots the ERDF units (92 Centres over the three year period analysed). In terms of the PLP model, the computed shadow price extreme values vary from 2.7 € to 15.7 €. And, as one would expect, marginal quality improvements tend to be more expensive as a network approaches 100% reliability.

INSERT FIGURE 2

7 **Conclusions**

In this paper we make use of the parametric distance function approach to identify the sources of technical inefficiency and the underlying technology in the energy distribution sector. Using panel data on 92 electricity distribution units operated by ERDF (*Electricité de France - Réseau Distribution*) over the 2003–2005 financial years, it appears that technical efficiency is positively correlated with the share of underground lines and with the age of capital. Moreover, the results show that shadow price of quality varies notably: from 2.7 € to 15.7 €, per customer interrupted among the operators. And, as one would expect, marginal quality improvements tend to be more expensive as a network approaches 100% reliability.

References

Battese G.E. and T.J. Coelli (1993), "A stochastic frontier production function incorporating a model for technical inefficiency effects", *Working Papers in Econometrics and Applied Statistics No 69*, Department of Econometrics, University of New England, Armidale.

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.

CEER (Council of European Energy Regulators), 2008, 4th Benchmarking Report on Quality of Electricity Supply, Brussels.

Coelli, T.J., Crespo, H., Plagnet, M.-A. and E. Romano (2007), "Incorporating quality of service in a benchmarking model: an application to French electricity distribution operators", CIRED, Vienna, 21-24 May 2007 Conference Proceedings.

Coelli, T.J. and S. Perelman (1996), "Eficiency measurement, multiple-output technologies and distance functions: With application to European railways", CREPP Working Paper 96/05, University of Liege.

Coelli, T.J., Estache, A., Perelman, S. and Trujillo, L. (2003), *A Primer on Efficiency Measurement for Utilities and Transport Regulators*, World Bank Publications.

Coelli, T.J. and D.S.P. Rao (2001), "Implicit Value Shares in Malmquist TFP Index Numbers", CEPA Working Papers, No. 4/2001, School of Economics, University of New England, Armidale.

Coelli, T.J., Rao, D.S.P., O'Donnell, C.J. and Battese, G.E. (2005), *An Introduction to Efficiency and Productivity Analysis*, 2nd Ed., Springer, New York.

Färe, R., Grosskopf, S., Lovell, K. C. A. and Yaisawarng, S. (1993), "Derivation of Shadow Prices for Undesirable Outputs: A Distance Function Approach", The Review of Economics and Statistics, 75(2), 374-380.

Giannakis, D., Jamasb, T., and Pollitt, M. (2005), "Benchmarking and Incentive Regulation of Quality of Service: An Application to the UK Electricity Distribution Utilities, *Energy Policy*, 33(17): 2256-2271.

Grosskopf, S., K. Hayes and J. Hirschberg (1995) "Fiscal Stress and the Production of Public Safety: A Distance Function Approach" *Journal of Public Economics* 57:277-296.

Growitsch, C., T. Jamasb and M. Pollitt (2005), "Quality of service, efficiency and scale in network industries: An analysis of European electricity distribution", IWH-DP 3/2005, IWH, Halle.

Fumagalli, E., Lo Schiavo, L. and F. Delestre (2007), *Service quality regulation in electricity distribution and retail*, Springer-Verlag Berlin Heidelberg.

Jamasb, T. and Pollitt, M. (2001) "Benchmarking and regulation: international electricity experience."

Utilities Policy, 9(3): 107-130.

Korhonen, P. and M. Syrjänen (2003), “Evaluation of Cost Efficiency in Finnish Electricity Distribution”, *Annals of Operations Research*, 121, 105-122.

Lawrence, D. and E. Diewert (2006), “Regulating Electricity Networks: The ABC of Setting X in New Zealand”, in Coelli, T.J. and D. Lawrence (eds.), *Performance Measurement and Regulation of Network Utilities*, Edward Elgar Publishing, Cheltenham, 207-237.

London Economics (1999) “Efficiency and benchmarking study of New South Wales distribution businesses” Report commissioned by the Independent Pricing and Regulatory Tribunal (IPART) of New South Wales, http://www.iprt.net/pdf/Rp13.pdf.

Morrison Paul, C. J. and R. Nehring (2005), “Product diversification, production systems, and economic performance in U.S. agricultural production”, *Journal of Econometrics*, 126, 525-548.

Newbery, D.M. (2000), *Privatization, Restructuring, and Regulation of Network Utilities*, Cambridge, MA: MIT Press.

O’Donnell, C. J. and T. J. Coelli (2005), “A Bayesian approach to imposing curvature on distance functions”, *Journal of Econometrics*, 126, 493-523.

.

Figure 1: Input oriented technical efficiency

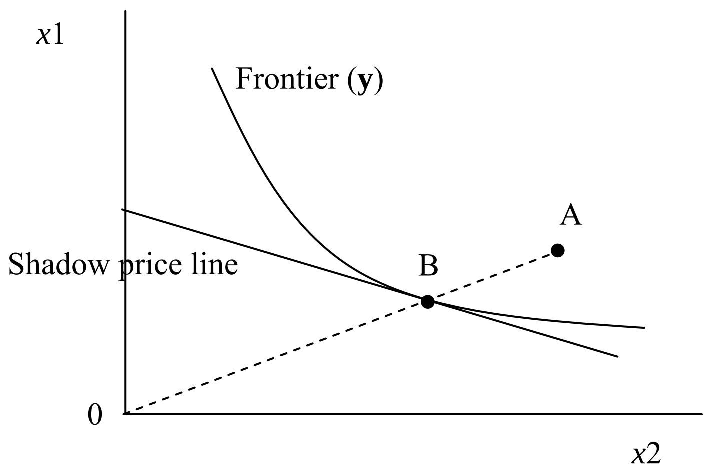

Figure 2: Quality shadow price

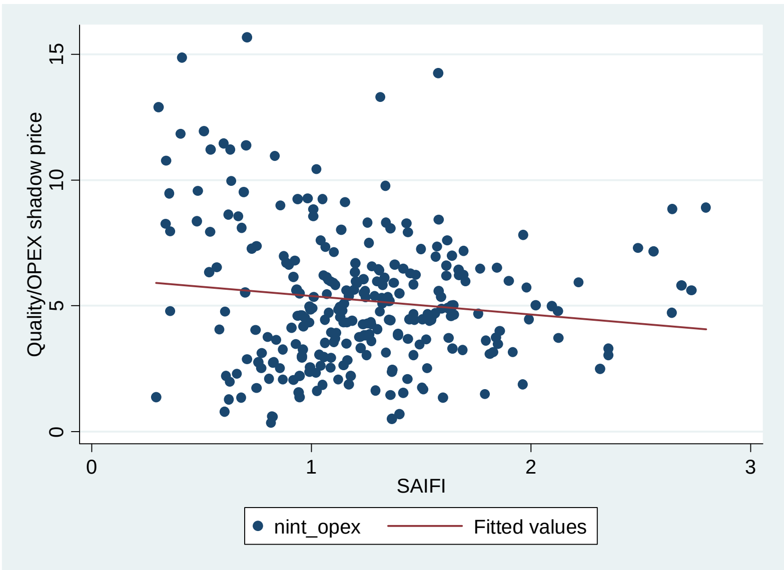

Table 1: Output and input variables. Descriptive statistics (n=276)

Variable	units	mean	std	min	max
Inputs					
Customers (CUST)	n	324,857	134,162	109,435	762,905
Surface (KM2)	km^2	5,532	3,125	129	13,871
Electricity (GWH)	GWh	3,557	1,477	1,001	7,976
Ouputs					
Operational expenditures (OPEX)	10^3 €	22,804	8,401	10,575	57,591
Capital (CAP)	10^3 €	636,134	220,018	247,464	1,250,115
Number of interruptions (NINT)	n	384,476	204,569	48,886	1,632,336
Environmental factors					
Underground lines (UNDERG)	ratio	0.39	0.19	0.13	0.88
Small towns (DENS)	ratio	0.44	0.23	0.00	0.86
Assets age proxy (AGE)	ratio	0.62	0.03	0.53	0.68
Customers growth (GROWTH)	%	1.97	1.81	-5.46	19.97
HV industrial capacity (HVCO)	ratio	0.24	0.06	0.14	0.42
Exceptional events:					
Frequency (EXNINT)	ratio	0.11	0.11	0.00	0.49
Minutes (EXMINT)	ratio	0.13	0.18	0.00	0.82

Table 2: Production characteristics

Customers' density quintile	Outputs ratios		Input / Output ratios		
	CUST / KM2	GWH / CUST	OPEX / CUST	CAP / KM2	NINT / CUST
Q1	23.0	11.0	91.4	63.9	1.53
Q2	36.3	12.2	77.6	84.5	1.27
Q3	55.5	12.2	73.4	118.4	1.33
Q4	88.2	11.9	67.7	163.9	1.05
Q5	1,223.0	10.6	58.9	1,384.0	0.97
Mean	288.6	13.1	73.7	366.7	1.23

Table 3: SFA and PLP coefficients

Explanatory variables	SFA		PLP	
	Coef.	(t-ratio)	Coef.	(t-ratio)
Intercept	0.154	(5.3) ***	-0.221	
$\ln(x_1)$ (OPEX)	0.501		0.501	
$\ln(x_2)$ (CAP)	0.457	(9.2) ***	0.456	
$\ln(x_3)$ (NINT)	0.041	(2.0) **	0.043	
$\ln(x_1).\ln(x_1)$	-0.706		-0.094	
$\ln(x_2).\ln(x_2)$	-0.829	(-3.1) ***	-0.211	
$\ln(x3).\ln(x_3)$	-0.136	(-1.7) *	-0.056	
$\ln(x_1).\ln(x_2)$	0.699		0.124	
$\ln(x_1).\ln(x_3)$	0.007		-0.030	
$\ln(x_2).\ln(x_3)$	0.129	(1.1)	0.087	
$\ln(y_1)$ (CUST)	-0.757	(-15.4) ***	-0.696	
$\ln(y_2)$ (KM2)	-0.101	(-5.3) ***	-0.107	
$\ln(y_3)$ (GWH)	-0.106	(-2.4) **	-0.227	
$\ln(y_1).\ln(y_1)$	0.949	(1.6) *	0.263	
$\ln(y_2).\ln(y_2)$	0.028	(2.1) **	-0.014	
$\ln(y_3).\ln(y_3)$	0.284	(0.5)	-0.392	
$\ln(y_1).\ln(y_2)$	0.004	(0.1)	-0.136	
$\ln(y_1).\ln(y_3)$	-0.658	(-1.2)	0.169	
$\ln(y_2).\ln(y_3)$	0.091	(1.6) *	0.027	
$\ln(x_1).\ln(y_1)$	-0.288		-0.083	
$\ln(x_1).\ln(y_2)$	0.055		0.058	
$\ln(x_1).\ln(y_3)$	0.243		-0.042	
$\ln(x_2).\ln(y_1)$	0.188	(0.6)	0.065	
$\ln(x_2).\ln(y_2)$	-0.052	(-1.2)	-0.072	
$\ln(x_2).\ln(y_3)$	-0.168	(-0.5)	0.048	
$\ln(x_3).\ln(y_1)$	0.100	(0.6)	0.017	
$\ln(x_3).\ln(y_2)$	-0.003	(-0.2)	0.014	
$\ln(x_3).\ln(y_3)$	-0.075	(-0.5)	-0.006	
Environmental factors				
Intercept	-0.266	(-1.8) *		
z_1 (D2004)	-0.005	(-0.3)		
z_2 (D2005)	-0.097	(-4.7) ***		
z_3 (UNDERG)	-0.226	(-6.2) ***		
z_4 (DENSE)	0.025	(0.6)		
z_5 (AGE)	0.629	(4.4) ***		
z_6 (GROWTH)	-0.001	(-0.2)		
z_7 (HVCON)	-0.027	(-1.2)		
z_8 (EXNINT)	0.007	(0.7)		
z_9 (EXMINT)	-0.007	(-1.2)		
σ	0.008	(11.5) ***		
γ	0.174	(6.0) ***		
LLF	286.0			

***, ** and * significant at 1%, 5% and 10% level, respectively. Underlined parameters are calculated by applying the homogeneity conditions. Variables $\ln(y_m)$ and $\ln(x_k)$ are expressed in deviations from sample mean values.

Table 4: SFA and PLP Technical Efficiency scores by quintiles

Quintiles	SFA			PLP		
	Customers' density	Quality (SAIFI)	Underground lines	Customer density	Quality (SAIFI)	Underground lines
Q1	0.803	0.853	0.783	0.664	0.755	0.688
Q2	0.843	0.870	0.843	0.807	0.783	0.805
Q3	0.858	0.879	0.872	0.822	0.832	0.833
Q4	0.915	0.907	0.925	0.884	0.871	0.881
Q5	0.981	0.960	0.982	0.861	0.878	0.866
All	0.897	0.897	0.897	0.828	0.828	0.828

Table 5: Distance function elasticities (PLP)

Customers' density quintiles	With respect to inputs "output shares"			With respect to outputs "input shares"			Scale elasticity
	Operational costs (OPEX)	Capital (CAP)	Quality (NINT)	Customers (CUST)	Surface (KM2)	Electricity distributed (GWH)	
Q1	0.385	0.584	0.032	-0.387	-0.140	-0.381	-0.908
Q2	0.447	0.524	0.029	-0.594	-0.100	-0.241	-0.934
Q3	0.490	0.473	0.038	-0.700	-0.086	-0.203	-0.988
Q4	0.538	0.426	0.037	-0.833	-0.064	-0.142	-1.034
Q5	0.670	0.259	0.076	-1.079	-0.098	-0.122	-1.262
Mean	0.533	0.424	0.046	-0.784	-0.092	-0.193	-1.037

Table 6: Shadow price ratios

Customers' density quintiles	Inputs			Outputs		
	Quality / OPEX	Quality / Capital	OPEX / Capital	Surface / Customers	GWH / Customers	Surface / GWH
Q1	4.9	97.1	20.5	9.0	106.0	0.09
Q2	3.7	94.9	26.4	6.6	38.9	0.19
Q3	4.2	125.6	30.7	7.0	27.5	0.30
Q4	4.2	140.5	35.3	7.6	18.9	1.07
Q5	7.5	613.5	75.7	117.9	13.4	29.29
All	5.1	254.1	42.1	37.5	32.5	7.977