ON ESTIMATING THE COST EFFICIENCY OF THE BRAZILIAN ELECTRICITY DISTRIBUTION UTILITIES USING DEA AND BAYESIAN SFA MODELS

Marcus Vinicius Pereira de Souza
PUC-RJ – Pontifícia Universidade Católica do Rio de Janeiro
Departamento de Engenharia Industrial
Rua Marquês de São Vicente 225 – Gávea 22451-041 – Rio de Janeiro – RJ
mvinic@engenharia.ufjf.br

Reinaldo Castro Souza
PUC-RJ – Pontifícia Universidade Católica do Rio de Janeiro
Departamento de Engenharia Elétrica
Rua Marquês de São Vicente 225 – Gávea 22451-041 – Rio de Janeiro – RJ
reinaldo@ele.puc-rio.br

Tara Keshar Nanda Baidya
PUC-RJ – Pontifícia Universidade Católica do Rio de Janeiro
Departamento de Engenharia Industrial
Rua Marquês de São Vicente 225 – Gávea 22451-041 – Rio de Janeiro – RJ
baidya@puc-rio.br

ABSTRACT
The purpose of this study is to evaluate the efficiency indices for 60 Brazilian electricity distribution utilities. These scores are obtained by DEA (Data Envelopment Analysis) and Bayesian Stochastic Frontier Analysis models, two techniques that can reduce the information asymmetry and improve the regulator’s skill to compare the performance of the utilities, a fundamental aspect in incentive regulation schemes. In addition, this paper also addresses the problem of identifying outliers and influential observations in deterministic nonparametric DEA models.

Keywords: Data envelopment analysis; Bayesian stochastic frontier analysis; Economic regulation; Outlier identifiers; Influential observations.

Introduction
In the Brazilian Electrical Sector (SEB, for short), the supply of energy tariffs is periodically revised within a period of 4 to 5 years, depending on the distributing utility contract. On the very year of the periodical revision, the tariffs are brought back to levels compatibles to its operational costs and to guarantee the adequate payback of the investments made by the utility, therefore, maintaining its Financial and Economical Equilibrium (EEF, for short). Over the period spanned between two revisions, the tariffs are annually readjusted by an index named IRT given by:

\[
IRT = \frac{VPA_1}{RA_0} + \frac{VPB_0(IGPM - X)}{RA_0} \tag{1}
\]

where, \( VPA_1 \) stands for the quantity related to the utility non-manageable costs (acquisition of energy and electrical sector taxes) at the date of the readjustment; \( RA_0 \) stands for the utility annual revenue estimated with the existing tariff (free of the ICMS tax) at the previous reference date IGPM (market prices index) and \( VPB_0 \) stands for the quantity related to the utility manageable costs (labor, third part contracts, depreciations, adequate payback of invested assets and working capital) on the previous reference date (\( VPB_0 = RA_0 - VPA_0 \)).
As shown in (1), the non-manageable costs (VPA) are entirely passed through to the final tariffs, while the amount related to the manageable costs (VPB) is updated using the IGPM index discounted by the $X$ factor. This factor applies only to the manageable costs and constitutes the way whereby the productivity gains of the utilities are shared with the final consumers due to the tariff reduction they introduce. The National Electrical Energy Agency (ANEEL) resolution 55/2004 defines the $X$ factor as the combination of the 3 components ($X_E, X_A$ and $X_C$), according to the expression below:

$$X = \{X_E + X_C\} \times \left[IGPM - X_A\right] + X_A$$

(2)

The component $X_A$ accounts for the effects of the application of the IPCA index (prices to consumer index) on the labor component of the VPB. The $X_C$ component is related to the consumer perceived quality of the utility service and the $X_E$ component accounts for the productivity expected gains of the utility due to the natural growth of its market. The latter is the most important and its definition is based on the discounted cash flow method of the forward looking type, in such a way to equal the present cash flow value of the utility during the period of the revision, added of its residual value, to the utility assets at the beginning of the revision period. In summary:

$$A_0 = \sum_{t=1}^{N} \frac{[RO_t (1 - X_E)]^{t-1} - T_t - OM_t - d_t (1 - g) + d_t - I_t}{(1 + r_{WACC})^t} + \frac{A_N}{(1 + r_{WACC})^N}$$

(3)

where, $N$ is the period, in years, between the two revisions; $A_0$ is the value of the utility assets on the date of the revision, $A_N$ is the utility assets value at the end of the revision period; $\beta$ stands for both; the income tax percentage and the compulsory social contribution of the utility applied to the utility liquid profit; $r_{WACC}$ is the average capital cost; $RO_t$ is the utility operational revenue; $T_t$ represents the various taxes (PIS/PASEP, COFINS and P&D); $OM_t$ are the operational and maintenance utility costs; $I_t$ is the amount corresponding to the investments realized and $d_t$ is the depreciation, all of them related to year $t$.

The quantities that form the cash flow in (3) are projected according to the criteria proposed by ANEEL, resolution 55/2004. As an example, the projected operational revenue is obtained as the product between the predicted marked and the average updated tariff; while the operational costs (operational plus maintenance, administration and management costs) are projected based on the costs of the “Reference Utility”, all related to the date of the tariff revision.

To avoid the complexity of the “Reference Utility” approach and in order to produce an objective way to obtain efficient operational costs, ANEEL envisages the possibility of using benchmarking techniques, among them, the efficient frontier method, as adopted by the same ANEEL to quantify the efficient operational costs of the Brazilian transmission lines utilities (ANEEL, 2007). The frontier is the geometric locus of the optimal production. The straightforward comparison of the frontier with the position of the utilities allows the quantification of the amount of improvement each utility should work on in order to improve its performance with respect to the others.

The international review conducted by Jasmab and Pollit (2000) shows that the most important benchmarking approaches used in regulation of the electricity services provided by utilities are based upon Data Envelopment Analysis (DEA, Cooper et al., 2000) and Stochastic Frontier Analysis (SFA, Kumbhakar and Lovell, 2000). As cited in Souza (2008), the first method is founded on linear programming, while the second is characterized by econometric models.

Studying cases of the SEB, authors such as Pessanha et al. (2004) and Sollero and Lins (2004) have used different DEA models to evaluate the efficiency of the Brazilian distributing utilities. On the other hand, Arcoverde et al. (2005) have also obtained efficient indices for the Brazilian distributing utilities using SFA models. Recently, Souza (2008) has proposed to gauge the cost efficiency using Bayesian Markov Chain Monte Carlo (MCMC) algorithm.
DEA and SFA approaches have distinct assumptions on their inner concept and present pros and cons, depending on the specific application. Therefore, there is no such statement as “the best” overall frontier analysis method.

In order to measure the efficiency (rather than inefficiency), and to make some interesting interpretations of efficiency across comparable firms, it is recommended to investigate efficiency indices obtained by several methods on the same data set, as carried out in the present work, where DEA and Bayesian SFA (BSFA hereafter) models are used to evaluate the operational costs efficiency of 60 Brazilian distributing utilities.

The paper is organized as follows. In the next section, it is discussed the basic concepts of the DEA and BSFA formulations. In addition, it is presented the Returns to Scale (RTS) question, the problem of detecting outliers, influential observations and Gibbs Sampler (MCMC) method. Section 3 comments on the results. Conclusions are given in Section 4. The appendix provides the main results obtained by DEA and BSFA methodologies summarized in Tables 2 and 3.

2. Methodology and Mathematical Models
2.1 The Deterministic DEA Approach

Data Envelopment Analysis is a mathematical programming based approach for assessing the comparative efficiency of the set of organisational units that perform similar tasks and for which inputs and outputs are available. It is meaningful to point out that in the DEA terminology, those entities are so-called Decision Making Units (DMUs).

The survey by Allen et al. (1997) reports that DEA was proposed originally by Farrell (1957) and developed, operationalised and popularised by Charnes et al. (1978). Ever since, this technique has been applied in a wide range of empirical work, such as education, banking, health care, public services, military units, electrical energy utilities, and others institutions. Zhu (2003) describes that one of the reasons for this argumentation could be that DEA has ability to measure the relative “technical efficiency” in a multiple inputs and multiple outputs situation, without the usual information on market prices.

In the framework here (DEA methodology), consider the case where there are $n$ DMUs to be evaluated. Each DMU $j$ ($j = 1, ..., n$) has consumed varying amounts of $m$ different inputs $x_j = [x_{1j}, ..., x_{mj}]^T \in \mathbb{R}^m$ to produce $s$ different outputs $y_j = [y_{1j}, ..., y_{sj}]^T \in \mathbb{R}^s$. A set of feasible combinations of input vectors and outputs vectors composes the Production Possibility Set $T$ (PPS, for short), defined by:

$$T = \{ [x, y] \in \mathbb{R}^{m+s} | x \text{ can produce } y \}$$

It is informative, here, to stress the study developed by Banker et al. (1984). In short, they postulated the following properties for the PPS, which are worthwhile:

- Postulate 1. Convexity;
- Postulate 2. Inefficiency Postulate;
- Postulate 3. Ray Unboundedness;

Subsequent to some algebraic manipulations under the above-mentioned four postulates, it is possible to show that the PPS $T$ is given by:

$$T = \{ [x, y] | x \geq X\lambda \geq 0 \}$$

where $X$ is the $(m \times n)$ input matrix, $Y$ is the $(s \times n)$ output matrix and $\lambda$ is a semipositive vector in $\mathbb{R}^n$.

If postulate 3 is removed from the properties of the PPS, it can be verified that:

$$T = \{ [x, y] | x \geq X\hat{\lambda}, y \geq Y\hat{\lambda}, \hat{\lambda} = 1, \lambda \geq 0 \}$$

where $\hat{\lambda}$ is the $(1 \times n)$ unit vector. A complete presentation of this demonstration, worth reading, can be found in Forni (2002).
Such results lead directly to two seminal DEA models. The first invokes the assumption of the Constant Returns-to-Scale (CRS) and convex technology, Charnes et al. (1978). On the other hand, the second assumes the hypothesis of Variable Returns to Scale (VRS), Banker et al. (1984).

In the following section, it is presented methods for measuring Return to Scale (RTS) of the technology.

2.2 Returns to Scale

As pointed out in Simar and Wilson (2002), it is very important to examine whether the underlying technology exhibits non-increasing, constant or non-decreasing RTS. Of course, large amount of literature has been developed on the problem of testing hypotheses regarding RTS. For example, Färe and Grosskopf (1985) suggested an approach for determining local RTS in the estimated frontier which involves comparing different DEA efficiency estimates obtained under the alternative assumptions of constant, variable, or non-increasing RTS, but did not provide a formal statistics test of returns to scale. On the other hand, Simar and Wilson (2002), again, discussed various statistics and presented bootstrap estimation procedures.

In some situations, it could be interesting to solve the RTS question by estimating total elasticity \( e \). Following Coelli et al. (1998), this estimate, certainly attractive from the point of view of simplicity, can be computed by using the partial elasticity estimates \( E_i \). However, it is easy to verify that this approach will fail in the very general setup of a multi output and multi-input scenario.

In terms of the partial elasticity estimates again, \( E_i \) is given by:

\[
e_i = \frac{\partial y_i}{\partial x_i} \cdot \frac{x_i}{y_i}
\]  

(7)

From its definition, the total elasticity \( e \) is expressed as:

\[
e = E_1 + E_2 + \ldots + E_n
\]  

(8)

Once the value of the total elasticity \( e \) is measured, immediately it is possible to identify the returns to scale type. Following Coelli et al. (1998), three possible cases are associated with (8):

- \( e = 1 \) ⇒ Constant Returns-to-Scale (CRS);
- \( e > 1 \) ⇒ Non-Decreasing Returns-to-Scale (NDRS);
- \( e < 1 \) ⇒ Non-Increasing Returns-to-Scale (NIRS).

In conformity with what is mentioned up to here, the next section focuses on how to find the feasible DEA model based on the resulting total elasticity.

2.3 DEA Models regarding Returns to Scale

As above, it is possible to determine the DEA best-practice frontier type through \( e \). In this context, let the CRS and VRS DEA models defined in (9) and (10) respectively:

\[
\min \left\{ \theta \mid y_0 \leq Y \lambda , \theta x_0 \geq X \lambda , \lambda \geq 0 \right\}
\]  

(9)

\[
\min \left\{ \theta \mid y_0 \leq Y \lambda , \theta x_0 \geq X \lambda , \lambda \hat{=} 1, \lambda \geq 0 \right\}
\]  

(10)

where \( \lambda \) is a \( n \times 1 \) row vector of weights to be computed, \( x_0 \) is a \( m \times 1 \) vector of inputs for DMU_0 and \( y_0 \) is a \( s \times 1 \) vector of outputs for DMU_0.

By inspection of (9) and (10), it is remarkable to notice that the VRS model (BCC model) differs from the CRS model (CCR model) only in the adjunction of the condition \( \lambda \hat{=} 1 \). Cooper et al. (2000) point out that this condition, together with the condition \( \lambda_j \geq 0, \forall j \), imposes a convexity condition on allowable ways in which the \( n \) DMUs may be combined.
Based on the appointed comments, it may be found in Zhu (2003) that if it is replaced \( \lambda_1 = \lambda \) with \( \lambda_1 \geq \lambda \), then it is obtained Non-Decreasing Returns-to-Scale (NDRS) model, alternatively, if it is replaced \( \lambda_1 = \lambda \) with \( \lambda_1 \leq \lambda \), then it is obtained Non-Increasing Returns-to-Scale (NIRS) model.

With regard to the interpretation of these models, it is straightforward: DEA minimize the relative efficiency index \( \theta \) of each DMU, comparing simultaneously all DMUs, subject to the constraints (remember that these constraints are equivalent to (5) and (6)).

Given the data, it is necessary to carry out an optimization for each of the \( n \) DMUs. Accordingly, a DMU is said to be fully efficient when \( \theta = \theta^* \) and, in this case, it is located on the efficiency frontier.

At this point another question arises: DEA models, by construction, are very sensitive to extreme values and to outliers. Even though Davies and Gather (1993) reasoned that the word outlier has never been given a precise definition, Simar (2003) defined an outlier as an atypical observation or a data point outlying the cloud of data points. This way, it is noteworthy that the outlier identification problem is of primary importance and it has been investigated extensively in the literature.

Besides this, it is important to stress that outliers can be considered influential observations. As stated by Dusansky and Wilson (1995), influential observations are those that result in a dramatic change in parameter estimates when they are removed from the data. For some interesting discussions about outliers and influential observations, see also Wilson (1993, 1995), Pastor et al. (1999), Forni (2002).

Herein, it is used to help detecting potential outlier the Wilson (1993) method. This technique generalizes the outlier measure proposed by Andrews and Pregibon (1978) to the case of multiple outputs and incorporates a convexity assumption. Nevertheless, as is seen from Wilson (1995), it becomes computationally infeasible as the number of observations and the dimension of the input-output space increases.

This discussion ends by assuming that these very rich results obtained will be extended in the BSFA context.

2.4 The Statistical Model

The stochastic frontier models (also known in literature as composed error models) were independently introduced by Meeusen and van den Broeck (1977), Aigner et al. (1977), Battese and Corra (1977) and have been used in numerous empirical applications. Some of the advantages of this approach are: a) identifying outliers in the sample; b) considering non manageable factors on the efficiency measurement.

Unfortunately, this method may be very restrictive because it imposes a functional form for technology.

This article uses a stochastic frontier model in Bayesian point of view. This technique allows to realize inference from data using probabilistic models for both quantities observed as for those not observed. Another feature of the BSFA framework is to enable the expert to include his previous knowledge in the model studied. For these reasons, Bayesian models are considered more flexible and thus, in most cases, they are not treatable analytically. To circumvent this problem, it is necessary to use simulation methods. The most used are the Markov Chain Monte Carlo methods (MCMC).

2.4.1 Bayesian Stochastic Cost Frontier

The econometric model with composed error for the estimation of the stochastic cost frontier can be mathematically expressed as:

\[
y_j = h(x_j; \beta) \exp\{v_j + u_j\}
\]  

Assuming that \( h(x_j; \beta) \) is linear on the logarithm, the following model is obtained after the application of a log transformation in (11):

\[
\ln y_j = \beta_0 + \sum_{i=1}^{m} \beta_i \ln x_{ji} + \sum_{i \leq k}^{m} \sum_{k=1}^{m} \beta_{ik} \ln x_{ji} \ln x_{jk} + v_j + u_j
\]
The equation (12) is called in literature as Translog function. When the crossed products are
onnull, there is a particular case called Cobb-Douglas function. With this information, the deterministic part of the frontier can be defined:

- \( \ln y_j \) - natural logarithm of the output of the \( j \)-th DMU (\( j = 1, \ldots, n \));
- \( \ln x_{ji} \) - natural logarithm of the \( i \)-th input of the \( j \)-th DMU (including the intercept);
- \( \beta = [\beta_0 \ \beta_1 \ \ldots \ \beta_m]^T \) - a vector of unknown parameters to be estimated.

In equation (12), the deviation between the observed production level and the determinist part of the frontier is given by the combination of two components:

- \( u_j \) an error that can only take non-negative values and captures the effect of the technical inefficiency, and
- \( v_j \) a symmetric error that captures any non manageable random shock. The hypothesis of symmetry of the distribution of \( v_j \) is supported by the fact that environmental favorable and unfavorable conditions are equally probable.

It is worthwhile to consider that \( v_j \) is independent and identically distributed (i.i.d, in short) with symmetric distribution, usually a Gaussian distribution, and that it is independent of \( u_j \). Taking into account the component \( u_j \) (\( u_j \geq 0 \)), this is not evident and thus can be specified by several ways. For example, Meeusen and van den Broeck (1977) used the exponential distribution, Aigner et al. (1977) recommended the Half-Normal distribution, Stevenson (1980) proposed the Truncated Normal distribution and finally, Greene (1990) suggested the Gamma distribution. More recently, Medrano and Migon (2004) used the lognormal distribution. The uncertainty related to the distribution of the random term \( u \) as well as the frontier function suggests the use of Bayesian inference techniques, as presented in pioneer works of van den Broeck et al. (1994) and Koop et al. (1995).

To this end, the sampling distribution is initially formulated. For example, considering the random term \( v_j \overset{iid}{\sim} N(0, \sigma^2) \), i.e., the Normal distribution with mean 0 and variance \( \sigma^2 \) and \( u_j \overset{iid}{\sim} \Gamma[1, \lambda^{-1}] \), i.e., \( u_j \overset{iid}{\sim} \exp[\lambda^{-1}] \), the joint distribution of \( y_j \) and \( u_j \), given \( x_j \) and the vector of parameters \( \psi \) (\( \psi = [\beta^T \ \sigma^2 \ \lambda^{-1}] \)) is given by:

\[
p(y_j, u_j | x_j, \psi) = N(y_j | h(x_j; \beta) + u_j, \sigma^2) \cdot \Gamma[u_j | 1, \lambda^{-1}]\] (13)

Integrating (13) with respect to \( u_j \), one arrives at the sampling distribution:

\[
p(y_j | x_j, \psi) = \lambda^{-1} \cdot \exp[-\lambda^{-1} \left( m_j + \frac{1}{2} \sigma^2 \lambda^{-1} \right)] \cdot \Phi \left( \frac{m_j}{\sigma} \right)\] (14)

where \( m_j = y_j - h(x_j; \beta) - \sigma^2 \lambda^{-1} \) and \( \Phi(\cdot) \) is the cumulative distribution function for a standard normal random variable.

To use the Bayesian approach, prior distributions are added to the parameters and, following the hierarchical modeling, posterior distributions are given. In principle, prior distribution of \( \psi \) may be any. However, it is usually non advisable to incorporate much subjective information on them and, in this case, appropriate prior specifications for the parameters need to be included. Here, consider the following prior distributions:

\[
\beta \sim N^+ \left(0, \sigma^2 \beta \right) \]

1 \( \Gamma(\cdot) \): Gamma function.
2 \( N^+ (\cdot, \cdot) \): Truncated Normal distribution.
According to Fernandez et al. (1997), it is essential that prior distribution $\sigma^{-2}$ is informative ($n_0 > 0$ and $c_0 > 0$) in order to ensure the existence of posterior distribution in stochastic frontier model with cross-section sample.

Following, in some cases, it is reasonable to identify similar characteristics among the companies evaluated and then, for including these informations in the model. This procedure can be performed specifying for each of DMUs, a vector $s_j$ consisting of $s_{jl}$ ($l = 1, \ldots, k$) exogenous variables. For these cases, Osiewalski and Steel (1998) proposed the following parameterization for the average efficiency:

$$\phi_j = \prod_{l=1}^{k} \phi_i^{-s_{jl}}$$

(17)

where $\phi_i > 0$ are the unknown parameters and, by construction, $s_{jl} = 1$. If $s_{jl}$ are dummy variables and $k > 1$, the distributions of $u_j$ may differ for different $j$. Thus, Koop et al. (1997) called this specification as Varying Efficiency Distribution model (VED, in short). If, $k = 1$, then $\phi_j = \phi_i^{-1}$ and all terms related to inefficiencies are independent samples of the same distribution. Again, according to Osiewalski and Steel (1998), this is a special case called Common Efficiency Distribution model (CED, in short).

Regarding to priori distribution of $k$ parameters of the efficiency distribution, Koop et al. (1997) suggested using $\theta_i \sim \Gamma(a_l,g_l)$ with $a_l = g_l = 1$ for $l = 2, \ldots, k$, $a_1 = 1$, and $g_1 = -\ln(r^*)$, where $r^* \in (0,1)$ is the hyperparameter to be determined. According to van den Broeck et al. (1994), in the CED model, $r^*$ can be interpreted as prior median efficiency. Proceeding this way, it could be ensured that the VED model is consistent with the CED model.

In agreement with the above, it is important to present posterior full conditional distributions of parameters involved in the model:

$$p(\sigma^{-2} | y_j, x_j, s_j, u_j, \beta, \phi) = p(\sigma^{-2} | y_j, x_j, u_j, \beta) \Gamma \left( \frac{n + n_0}{2}, \frac{c_0 + \sum_j (y_j - h(x_j; \beta) - u_j)^2}{2} \right)$$

(18)

$$p(\beta | y_j, x_j, s_j, u_j, \sigma^{-2}, \phi) = p(\beta | y_j, x_j, u_j, \sigma^{-2}) \sim N^+ \left( \beta | 0, \sigma^{-2} \right) \exp \left( -\frac{1}{2} \sigma^{-2} \sum_j (y_j - h(x_j; \beta) - u_j)^2 \right)$$

(19)

The posterior full conditional distribution of $\phi_j$ ($l = 1, \ldots, k$) presents the following general form:

$$p(\phi_l | y_j, x_j, s_j, u_j, \beta, \sigma^{-2}, \phi_{(-l)}) = p(\phi_l | s_j, \phi_{(-l)}) \propto \exp \left( -\phi_l \sum_j u_j D_{jl} \right) \chi(\phi_l^2 + \sum_j s_{jl} g_l)$$

(20)

where:

$$D_{jl} = \prod_{j=1}^{k} \phi_j^{s_{jl}}$$

(21)

For $l = 1, \ldots, k$ ($D_{jl} = 1$ for $k = 1$) and $\phi_{(-l)}$ denotes $\phi$ without its $l$-th element.

With regard to inefficiencies, it can be shown that they are distributed as a Truncated Normal distribution.
\[
p_{u_j|y_j, x_j, s_j, \beta, \sigma^2, \phi} = \left[ \phi \left( \frac{h(x_j; \beta) - y_j - \lambda \sigma^2}{\sigma} \right) \right]^{-1} \times N(u_j|h(x_j; \beta) - y_j - \lambda \sigma^2, \sigma^2) \tag{22}
\]

As the posterior full conditional distribution for \( u \) is known, Gibbs sampler could be used to generate observations of the joint posterior density. These observations could be used to make inferences about the unknown quantities of interest. It is worth remembering that the technical efficiency of each DMU is determined making \( r_j = \exp(-u_j) \).

### 2.4.2 The Gibbs Sampler (MCMC) Algorithm

According to Gamerman (1997), the Gibbs sampler was originally designed within the context of reconstruction of images and belongs to a large class of stochastic simulation schemes that use Markov chains. Although it is a special case of Metropolis-Hastings algorithm, it has two features, namely:

- All the points generated are accepted;
- There is a need to know the full conditional distribution.

The full conditional distribution is the distribution of the \( i \)-th component of the vector of parameters \( \psi \), conditional on all other components.

Again referring to Gamerman (1997), the Gibbs sampler is essentially a sampling iterative scheme of a Markov chain, whose transition kernel is formed by the full conditional distributions.

To describe this algorithm, suppose that the distribution of interest is \( p(\psi) \), where \( \psi = (\psi_1, \ldots, \psi_d) \). Each of the components \( \psi_i \) can be a scalar, a vector or a matrix. It should be emphasized that the distribution \( P \) does not, necessarily, need to be an a posteriori distribution. The implementation of the algorithm is done according to the following steps (Gamerman (1997)):

1. initialize the iteration counter of the chain \( t = 1 \) and set initial values \( \psi^{(0)} = (\psi_1^{(0)}, \ldots, \psi_d^{(0)}) \);
2. obtain a new value \( \psi^{(t)} = (\psi_1^{(t)}, \ldots, \psi_d^{(t)}) \) from \( \psi^{(t-1)} \) through successive generation of values:
   
   \[
   \psi_1^{(t)} \sim p(\psi_1|\psi_2^{(t-1)}, \ldots, \psi_d^{(t-1)})
   \]
   
   \[
   \psi_2^{(t)} \sim p(\psi_2|\psi_1^{(t)}, \psi_3^{(t-1)}, \ldots, \psi_d^{(t-1)})
   \]
   
   \[
   \vdots
   \]
   
   \[
   \psi_d^{(t)} \sim p(\psi_d|\psi_1^{(t)}, \ldots, \psi_{d-1}^{(t)})
   \]
3. change counter \( t \) to \( t+1 \) and return to step (ii) until convergence is reached.

Thus, each iteration is completed after \( d \) movements along the coordinated axes of components of \( \psi \). After convergence, the resulting values form a sample of \( p(\psi) \). Ehlers (2005) emphasizes that even in problems involving large dimensions, univariate or block simulations are used which, in general, is a computational advantage. This has contributed significantly to the implementation of this methodology, especially in applied econometrics area with Bayesian emphasis.

### 3. Experimental Results and Interpretation

To evaluate the efficiency, the utilities have been characterized by the 4 indicators marked in Table 1. The products are the cost drivers of the operational costs. The amount of energy distributed (MWh) is a proxy of the total production, the number of consumer units (NC) is a proxy for the quantity of services provided and the grid extension attribute (KM) reflects the spread out of consumers within the concession area, an important element of the operational costs.
By now, it is useful to start for identifying the outliers among the utilities. To do so, the Wilson (1993) method was used and the information on 60 Brazilian utilities was processed by the FEAR 1.11 (a software library that can be linked to the general-purpose statistical package R). Then the following utilities have been considered outliers: CEEE, CELPA, PIRATININGA, BANDEIRANTES, CEB, CELES, CELG, CEMAT, CEMIG, COPEL, CPFL, ELETROPAULO, ENERSUL, LIGHT. It can be observed that this technique has classified the utilities with the largest markets, with geographical concentration and a strong industrial share of participation. For instance: BANDEIRANTES, CEMIG, COPEL, CPFL, ELETROPAULO, ENERSUL, LIGHT.

Results concerning to the measurement of efficiency were obtained by NIRS DEA models because the total elasticity \( e \) is less than 1 (report to section 2.2 and 2.3).

### Table 1 – Input and Outputs variables.

<table>
<thead>
<tr>
<th>Type</th>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input (DEA) or dependent (BSFA)</td>
<td>OPEX</td>
<td>Operational Expenditure (R$ 1.000).</td>
</tr>
<tr>
<td>Output (DEA) or independent (BSFA)</td>
<td>MWh</td>
<td>Energy distributed.</td>
</tr>
<tr>
<td></td>
<td>NC</td>
<td>Units consumers.</td>
</tr>
<tr>
<td></td>
<td>KM</td>
<td>Network distribution length.</td>
</tr>
</tbody>
</table>

The scores for each of the 60 DMUs are exhibited in Table 2 in the appendix. They were calculated using the DEA Excel Solver developed by Zhu (2003).

By analyzing the scores obtained by R1 in Table 2, it can be observed that nine companies are on the best-practice frontier. Note also that seven (PIRATININGA, BANDEIRANTES, CEMIG, COPEL, CPFL, ELETROPAULO and ENERSUL) were labeled as outliers. Given these results, it is meaningful to emphasize that these seven DMUs will be considered influential observations. A good strategy to verify this fact is to carry out another NIRS model on both the set of outlier DMUs (14 companies) and the remaining DMUs.

Focusing attention on the results listed in Table 2, it can be ascertained that the performance of inefficient DMUs (see R1), which has as benchmarks the outlier companies, has changed so much (refer to R2). That is, the average range of efficiency variation to this set of DMUs is approximately 10%. This case can be mathematically expressed as:

\[
\Delta \text{Eff} = \frac{1}{31} \sum_{j=1}^{31} \frac{R_{2,j} - R_{1,j}}{R_{1,j}}
\]

\( j = 2, 5, 6, 7, 8, 10, 13, 14, 17, 18, 20, 21, 24, 25, 28, 31, 33, 34, 35, 40, 41, 42, 43, 46, 52, 54, 55, 56, 57, 59, 60. \)

Continuing further, for others inefficient DMUs, it is approximately 1%. Mathematically:

\[
\Delta \text{Eff} = \frac{1}{20} \sum_{j=1}^{20} \frac{R_{2,j} - R_{1,j}}{R_{1,j}}
\]

\( j = 3, 4, 9, 11, 16, 23, 32, 36, 37, 38, 39, 44, 45, 47, 48, 49, 50, 51, 53, 58. \)

With respect to econometric methodology, it is important to attribute a specification for the costs frontier. To this end, a Cobb-Douglas functional form was adopted, which is defined by:

\[
\ln \text{OPEX} = \beta_0 + \beta_1 \ln \text{MWh}_j + \beta_2 \ln \text{NC}_j + \beta_3 \ln \text{KM}_j + v_j + u_j
\]

As mentioned in section 2.4.1, it is useful that the expert incorporates information on companies to the model. Accordingly, by inspection of R2 in Table 2, is possible to obtain the following: (i) Efficient NIRS DEA utilities; (ii) Prior median efficiency.

The first consideration suggests the use of a dummy variable while the second provides \( r^* = 0.620 \) (approximate value of the prior median efficiency evaluated in Table 2, in the column headed “Adjusted input oriented NIRS Efficiencies (R2)”).

---

1 The FEAR package is available at: http://www.economics.clemson.edu/faculty/wilson/Software/FEAR
The Bayesian model is carried out using the free software WinBUGS (Bayesian inference Using Gibbs Sampling for Windows) that can be downloaded at www.mrc-bsu.cam.ac.uk/bugs/Welcome.htm.

In this context, the chain was run with a burn-in of 20,000 iterations with 50,000 retained draws and a thinning to every 7-th draw. WinBUGS has a number of tools for reporting the posterior distribution. A simple summary (see Table 3 in the appendix) can be generated showing posterior mean, median and standard deviation with a 95% posterior credible interval. Referring to estimated coefficients, it is worth reporting that they are significant. To examine the question of analysis of convergence of parameters, this was verified using serial autocorrelation graphs.

Finally, the Pearson correlation coefficients as well as the Spearman rank-order correlation coefficients computed between the model estimates (R1, R2 and R3) are statistically significant at the 5% level and ranges from 83 to 96 percent.

4. Conclusions

The measurement of efficiency obtained by the DEA and Bayesian SFA model should express the reduction in operational costs. In accordance with that has been already exposed, the potential reduction of the operational costs for the \( j \)-th utility, i.e., the operational cost recognized by the regulator is equal to \( \text{OPEX}_j \times (1 - \theta_j) \).

For the next tariff revision cycles, the Brazilian regulator ANEEL has signalized with the possibility of using DEA and SFA models in the estimation of the efficient operational costs, an important element in the determination of the X-factor of the utilities. The two approaches use different assumptions; the DEA is deterministic and deviations with respect to the efficient frontier are assumed to be solely due to the utilities inefficiency, whereas the SFA has a stochastic nature and provides estimates of the efficiency, free of the uncontrollable impacts of random factors that affect the DMUs.

In this work, it can be ascertained that the conjoint analysis of DEA and Stochastic Frontier in the Bayesian approach is fundamental. Indeed, this is demonstrated through easy incorporation of prior ideas and formal treatment of parameter and model uncertainty.

Appendix

Table 2 – Efficiency scores (\( \theta_j \)).
<table>
<thead>
<tr>
<th>DMU Name</th>
<th>Local oriented DMU Efficiencies (R2)</th>
<th>Adjusted local oriented DMU Efficiencies (R2)</th>
<th>Benchmarks</th>
<th>Adjusted local oriented DMU Efficiencies (R2)</th>
<th>Benchmarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMU 1</td>
<td>0.977</td>
<td>0.000</td>
<td>0.950</td>
<td>0.000</td>
<td>0.950</td>
</tr>
<tr>
<td>DMU 2</td>
<td>0.920</td>
<td>0.050</td>
<td>0.880</td>
<td>0.050</td>
<td>0.880</td>
</tr>
<tr>
<td>DMU 3</td>
<td>0.958</td>
<td>0.020</td>
<td>0.918</td>
<td>0.020</td>
<td>0.918</td>
</tr>
<tr>
<td>DMU 4</td>
<td>0.879</td>
<td>0.100</td>
<td>0.819</td>
<td>0.100</td>
<td>0.819</td>
</tr>
<tr>
<td>DMU 5</td>
<td>0.975</td>
<td>0.000</td>
<td>0.945</td>
<td>0.000</td>
<td>0.945</td>
</tr>
<tr>
<td>DMU 6</td>
<td>0.920</td>
<td>0.050</td>
<td>0.880</td>
<td>0.050</td>
<td>0.880</td>
</tr>
<tr>
<td>DMU 7</td>
<td>0.958</td>
<td>0.020</td>
<td>0.918</td>
<td>0.020</td>
<td>0.918</td>
</tr>
<tr>
<td>DMU 8</td>
<td>0.879</td>
<td>0.100</td>
<td>0.819</td>
<td>0.100</td>
<td>0.819</td>
</tr>
<tr>
<td>DMU 9</td>
<td>0.977</td>
<td>0.000</td>
<td>0.950</td>
<td>0.000</td>
<td>0.950</td>
</tr>
<tr>
<td>DMU 10</td>
<td>0.920</td>
<td>0.050</td>
<td>0.880</td>
<td>0.050</td>
<td>0.880</td>
</tr>
<tr>
<td>DMU 11</td>
<td>0.958</td>
<td>0.020</td>
<td>0.918</td>
<td>0.020</td>
<td>0.918</td>
</tr>
<tr>
<td>DMU 12</td>
<td>0.879</td>
<td>0.100</td>
<td>0.819</td>
<td>0.100</td>
<td>0.819</td>
</tr>
</tbody>
</table>

Table 3 – Results of efficiencies obtained by BSFA.

References


ANEEL Resolução Normativa nº 55/2004, 5 de abril de 2004

ANEEL Nota Técnica nº 166/2006 – SRE/ANEEL, 19 de maio de 2006a

ANEEL Nota Técnica nº 262/2006 – SER/SFF/SRD/SFE/SC/ANEEL, 19 de outubro de 2006b


