The efficiency of Brazilian electricity distributors during 2004 – 2009. An application using DEA corrected by environmental and stochastic factors.

Fernando Damonte[1], Quantum, fdamonte@quantumamerica.com

Mariana De Santis[2], Quantum, mdsantis@quantumamerica.com

Sanford Berg[3], University of Florida, sberg@ufl.edu

Abstract

In this study, efficiency scores of Brazilian electricity distributors for the period 2004 - 2009 are estimated. With a database from the Brazilian regulator, the analysis uses a three-stage methodology by Fried et al (2002). At an initial stage DEA is applied, using electricity sales and customers as outputs and capital costs and OPEX as inputs. In a second stage, the inputs' efficient slacks are obtained, after applying a SFA model to correct for the effects of environmental and random variables. Finally, each of the efficiency scores is adjusted by applying a common operating environment to each company. Alternative models were also specified, resulting in greater estimates of efficiency after adjusting for stochastic and environmental factors. These models do not present sharp changes over the period. This methodology overcomes limitations associated with the one employed recently by the Brazilian regulator: a DEA model adjusting for environment variables.

***Keywords**: efficiency, DEA, Multistage, Stochastic Frontier Analysis, electricity distribution, Brazil*

1. Introduction

This study estimates the efficiency of Brazilian electricity distribution companies during 2004 to 2009. A semi-parametric methodology that incorporates the impact of both environmental variables and statistical noise is applied. DEA is initially employed, supplemented with SFA methodology, following Fried et al (2002). This methodology adjusts the amounts of inputs by incorporating environmental variables and statistical noise from a stochastic frontier approach. After bringing all companies into a common scenario, the third step involves recalculating the efficiency scores for the DEA model. Unlike other methodologies, the amount of each input is adjusted instead of the efficiency score, allowing environmental variables to differentially affect each of the inputs. When a utility faces unique disadvantages (captured by the regulator's "complexity index"), this approach

[1] Expert consultant at Quantum America Corp.

[2] Expert consultant at America Corp.

[3] Distinguished Service Professor—Department of Economics, Warrington College of Business Administration. He is also Director of Water Studies, Public Utility Research Center, University of Florida.

The authors are grateful for comments and helpful suggestions by Rui Marques and Chunrong Ai. The authors are solely responsible for the content.

recognizes that additional specific inputs (such as labor) will need to be utilized to produce the same output as a firm serving a geographic franchise area with a lower socioeconomic complexity index (reflecting levels of violence, economic deprivation, and other factors). This methodology is particularly advantageous when estimates of efficiency utilize more than one input, such as total expenditures, operating expenditures and capital expenditures.

Studies of electricity distribution efficiency have attracted increasing attention since the 90s, when a large number of utility regulators around the world, as part of reforms in the industry structure and tariff setting, replaced the traditional system of rate of return regulation with incentive-based schemes. The central objective of incentive regulation is to promote improvements in the efficiency of regulated industries in the absence of market mechanisms. So understanding trends, establishing baselines, and identifying best practice have become hallmarks of regulatory activity.

Overall, benchmarking can be defined as the comparison of performance indicators of real companies with a "benchmark" or reference. In a benchmarking study, companies are considered to be production entities that transform inputs into outputs. The crucial problem is that the regulator does not generally have accurate and sufficient information to determine the optimal amount of inputs used by a regulated company to obtain a certain level of product. If the inputs are considered in monetary terms, i.e. dealing with production costs, the regulator needs a criterion to determine whether the cost incurred by a particular company is appropriate or if those costs can be decreased without sacrificing some dimension of output. One technique for overcoming the information asymmetry is to use production data and costs of companies in the same industry to infer the level of costs or the achievable goal for a particular company. This information is then used in determining the rates, with firms far from the frontier penalized through the approval of lower rates. Presumably, the relatively inefficient utility can reduce its inputs without sacrificing output.

The practice of using benchmarking studies for setting prices is widespread. Regulation is based increasingly on the results of benchmarking studies since techniques have improved over time; these studies allow the regulator to measure the relative operating efficiency for each company. From these efficiency studies, the targets for companies can be established with greater precision. In the Brazilian electricity distribution sector, benchmarking techniques are being used by the regulatory agency to validate results from different models and to identify efficient operating costs for each company. Although the successive methodologies applied by the Brazilian regulator in the determining allowable prices (and associated revenues) have been improving, it is still necessary to examine the effects of stochastic factors and environmental features beyond managerial control on the cost frontier. In addition, the potential effect of the firm size on efficiency warrants attention.

This study estimates the technical efficiency of Brazilian electricity distribution companies, considering stochastic effects, non-managerial variables and total costs. The results suggest that the impact of random effects and the trade-off between OPEX and CAPEX should be taken into account in efficiency analysis. This paper is organized as follows. Section 2 reviews the main studies by academic and regulatory authors regarding efficiency in the Brazilian electricity distribution sector. In Section 3 the technology of electricity distribution is analyzed in order to represent it properly in the models estimated. Then Section 4 presents the methodology used in the analysis, evaluates its suitability providing

efficiency scores, and describes the data and variables used to estimate the model. The empirical results and analyses are presented in Section 5. Finally, Section 6 summarizes the main findings of this study.

2. Antecedents

The literature about estimations electricity distribution sector efficiency is large and growing, especially with regards to European countries. However, the evidence for Latin American countries is limited due to data availability issues and the relatively recent introduction of benchmarking techniques for tariff setting. Here, only studies about Brazilian electricity firms are reviewed, including academic papers and regulatory reports published by the regulatory agency (ANEEL).

Mota (2004) analyzes the performance of the electricity distribution companies in Brazil in relation to their peers in USA in 1994 and 2000, with the objective of evaluating the impact of privatization on the efficiency of the sector. The author uses DEA to estimate efficiency scores in each year and decompose 1994 – 2004 Malmquist indexes. On the other hand, she estimates SFA models by specifying different input-oriented distance functions. Mota reviews benchmarking studies in the electricity sector and concludes that there is no uniform approach to choosing variables reflecting the output and inputs. However, since regulators are interested in evaluating the efficiency of production costs, most regulatory studies use monetary, instead of physical variables as indicators of inputs. Commonly used monetary variables are operating costs and capital costs.

Due to the reason mentioned above, in the study by Mota both operating expenses (OPEX) and total costs (TOTEX) were used as proxies for the inputs. The author presents six different specifications, including the energy sold, the number of customers and network length as outputs. In addition, in some models, environmental variables such as peak demand, customer density and the proportion of residential customers were included.

The data used in the estimates come from 72 companies owned by American investors (investor-owned utilities) and 14 Brazilian companies. Variables are referred to the years 1994 and 2000. In terms of data sources, information for U.S. companies was taken from FERC reports and data for the Brazilian companies from annual reports by companies and fieldwork.

This early study concluded that while higher efficiency estimates are obtained for Brazilian companies in relation to the American ones with both SFA and DEA approaches, the difference is not statistically significant. It cannot be concluded that in each time period Brazilian distributors have better performance than U.S. companies. In examining the impact of privatization on efficiency electricity distribution sector in Brazil, the author notes that it is positive but insignificant when it is estimated using OPEX models, while it is negative and statistically significant when models that include TOTEX are considered. These results suggest that the process of investment in the sector driven by privatization led to a replacement of the physical labor by capital (or to network expansion), which is why all estimated models should include both the cost of labor and the cost of capital. Furthermore, the results of stochastic frontier models using panel data indicate a significant technology gap favoring Brazilian companies, whether operating costs or total costs are

considered. However, the estimated rate of technological change was greater for American companies.

In a subsequent paper, Ramos-Real et al (2009) estimate the evolution of productivity in the electricity distribution industry in Brazil during 1998 to 2005, with the objective of evaluating the impact of the sector reforms which took place in this period. The authors employ a panel of 17 distributors. The data were taken from the Brazilian Association of Electricity Distribution Companies ABRADEE reports and complemented with information from the annual reports of the companies under study. The 17 companies distributed about 55% of Brazil's electricity consumption in 2005. In terms of methodology, the authors employ the Malmquist TFP index to analyze the change in the productivity of the sector by using DEA estimates. The study suggests that in a DEA context, estimates of Malmquist productivity index, as well as its sources of growth is given by ratios of distance functions values for technologies with constant and variable returns to scale. The sources of productivity growth are technical efficiency and technical change, which can be decomposed into pure efficiency and scale efficiency. Variables representing distributors' outputs are energy delivered and the number of customers, while the inputs considered are the number of employees, the length of the network (as proxy of physical capital) and total losses (both theft and line losses). Finally, the size of the service area was incorporated as an environmental variable beyond managerial control.

The results show that during the period 1998-2005 the rate of total factor productivity increased at a 1.3% annually, on average. This was decomposed into an increase in technical change of 2.1% per year, offset by a decrease in annual average technical efficiency of 0.8%. The authors conclude that the variation in productivity is attributed to technical change, which implies a shift of the frontier, attributed to technological innovations. However, the results of pure technical efficiency indicate that the firms under study did not improve their performance, suggesting that reforms in the regulation of the sector did not produce the expected results (moving firms closer to the expanding frontier). Furthermore, when analyzing rates from year to year, it is found that between 2001 and 2002 and between 2002 and 2003, productivity declined, probably attributable to the energy shortage experienced by the country in 2001. In addition, Ramos-Real found a fall in pure efficiency and scale efficiency, respectively, while in 2004 and 2005 a rebound in technical efficiency was observed.

In a later work, using the same data, Tovar et al (2011) analyze the evolution of total factor productivity (TFP) in the Brazilian electricity distribution sector by decomposing TFP into technical efficiency, technical change and scale effect. The authors find the productivity development for the sampled companies during the period 1998-2003 was attributable to the frontier shift; that is to say, technical change was mainly due to technological innovations. With regards to pure technical efficiency change, their study finds that distribution utilities are moving away from the efficiency frontier, and have not improved their performance—regardless firm size. The authors conclude that the scale effect is relevant, suggesting that mergers might improve the sector performance.

More recently, Silva (2011) analyzes the correspondence between the OPEX efficiency scores obtained by the Brazilian regulator by comparing the regulatory model company (an engineering approach to efficiency) and scores derived from alternative quantitative

methods. The author compares the efficient OPEX determined by the regulatory agency following the reference business model for 2003 to efficient OPEX estimated according to a SFA model and a multi-stage DEA methodology. Estimations are based on an unbalanced panel of 52 companies, responsible for 99.47% of the total electricity delivered in 2003; with the data were collected for the 1998 to 2003 time period. Silva finds the distributors' ranking from the model company and that from SFA are not significantly correlated, with only four of the same companies appearing on top ten lists from both methodologies and only two appearing on both the bottom ten lists. In general, companies are more efficient according to the SFA than the engineering (model company) approach, because the latter is not based on the actual network, but on a hypothetical "optimized" network—starting from scratch. Utilizing the efficiency score estimates from the engineering approach in the calculation of revenue requirements for the tariff revision is inappropriate according to Silva. He corrects his DEA estimates to incorporate the impact of environmental variables by estimating a three-stage DEA although without estimating a SFA to compute the efficient slacks. The results show significant relative efficiencies compared with the use of the model company technique. By overestimating the performance of some distributors, these utilities received unjustified benefits during the tariff calculation.

The evolution of benchmarking methodologies is reflected in the three most recent tariff reviews. In the first regulatory review that occurred in 2003, ANEEL continued the deep reform of the electricity sector that had begun in 1995 with the privatization of two large regional distributors. The model firm (engineering approach) was used by developing a detailed mathematical representation of the activities and processes implemented by each electricity distribution company. In the second tariff revision in 2007, the model company was again applied, although this time the results were compared with those of alternative benchmarking approaches to determine if the model reflected captured key features of the sector, such as average level of efficiency. In fact, the regulator used benchmarking techniques in three ways: i) to analyze the global consistency of predicted operating costs obtained by applying the model company method; ii)to determine approved non-technical losses by using delinquency rates (robbery homicides per 100, 000 inhabitants) as a proxy for the society's propensity to fraud; and iii) to determine the quality of service. With its analysis of global consistency, ANEEL attempted to validate the reasonableness of the costs of the model company, obtained by "Top-Down" with an alternative methodology. These analyses were not made explicit in the discussions with the agents involved in the process. However, the agency committed to explaining the global consistency of analyzes to be used in the third review.

In 2011, the third tariff review of the Brazilian electricity distribution sector took place. This time, ANEEL applied a combination of parametric and nonparametric methods to estimate efficient regulatory OPEX targets to be achieved by the companies at the end of the three year cycle. This methodology involved a significant departure from past approaches. In the latest review, which began in 2010, operating costs of distributors were calculated by updating the operating costs obtained in second cycle: the adjustments reflected productivity changes and inflation. However, in addition, ANEEL introduced the use of benchmarking techniques based on DEA and Corrected Ordinary Least Squares (COLS) to determine the efficient operating costs to be achieved by each company at the end of the cycle, in 2014.

ANEEL used data from all distribution companies during the period 2003-2009 and applied a two-stage frontier estimation methodology. In the first stage, the efficiency was estimated by two methods: (1) Non Decreasing Returns to Scale (NDRS) input-oriented DEA using OPEX as a single input and distributed energy, number of customers and network length as three outputs; and (2) a linear function in logarithms OPEX was estimated, considering the three outputs mentioned as explanatory variables, applying the COLS model. The average efficiency scores resulting from both models were adjusted for the effect of environmental variables: average wage paid by each firm, precipitation level, social environment and consumer density. For this last correction, three alternative methods were employed[4] and their respective results were then averaged. These methods consisted in (1) running a regression of the efficiency scores against environmental variables applying a Tobit model, (2) applying the methodology suggested by Simar and Wilson (2007), and (3) utilizing the methodology (that included contextual variables) from a study by Banker and Natarajan (2008). The three procedures were applied separately to large (more than one TWh/year) and small distributors.

Thus, many estimates of Brazilian electricity distribution companies' efficiency have been conducted using different approaches. However, the typical issues that arise include accounting for stochastic effects, the influence of non-managerial variables, multi-collinearity among outputs, and insufficient precision (reflecting model specification) in representing the nature of the electricity distribution business. These issues have been only partially resolved by the above-mentioned studies. The current study attempts to address them all by including random effects and environmental variables; the application of DEA allows the use of multi-collinear outputs and breaks down the output variables into more granular data to gain precision in the estimation.

2. The technology of electricity distribution

The specification of empirical models for the estimation of the electricity distributors` efficiency requires a detailed analysis of variables that should be considered as outputs and inputs, as well as those context variables which affect distribution costs. We begin with a classic study by Neuberg (1977), which analyzes the technology of electricity distribution using a cost function. That model was used in the estimation of economies of scale and efficiency of U.S. distributors in 1972. Neuberg concludes, after reviewing engineering studies of the distribution business, that the number of customers can be considered as an output. In addition, he shows that the network length, the amount of power sold and surface area should also be considered as "cost drivers". Neuberg argues that the number of customers, network size, energy sold (kWh), and the area served cannot be considered to be products resulting from a joint production process, because they cannot be sold separately in individual markets. When being considered as inputs, they should not be introduced into the cost function, since cost depends on the amount of output and input prices, but not of the physical quantities of inputs.

[4] See Technical Note Number 294/2011, SRE/ANEEL

Alternatively he estimated a production function in which the extension of the network, area and energy sold as inputs entered, using the primal approach. This approach estimated the production function instead of the cost function to analyze the properties of the technology and the efficiency of the firm. However, Neuberg noted the endogeneity of these inputs. While an electricity distributor cannot ignore the network when providing the service, the network is not endogenous, but can, like energy sold and the concession area, be regarded as given and fixed. The obvious question is: if the factors are neither strictly outputs nor inputs, what are they? Neuberg called them "non-classical variables", i.e. variables that can be considered as exogenous to the firm and as in the case of output levels, having an impact on costs. The non-classical variables reflect the particular physical characteristics of distributors which affect costs, such as customer density. Lower density can justify higher costs for a firm that, serving the same number of customers as another, must invest in a network that covers a larger area.

In a subsequent study on the efficiency of electricity companies in England and Wales, Burns and Weyman-Jones (1996) explain the operating costs in terms of the energy delivered and peak demand, controlling for exogenous variables that reflect the environmental characteristics in which companies operate: network length, power transformers, market structure, and customer density. The authors note that the inclusion of exogenous variables allows comparability between firms with different levels of business efficiency according to Shleifer's (1985) yardstick competition model. Burns and Weyman-Jones examine the choice of output variables choice using the criteria of separately offered and priced items.

The current study views distribution companies as multiproduct firms that could provide their services separately, outsourcing these services or competing against other firms under a framework developed by the regulator. An example of this is the introduction of retail competition in some countries. For the electricity distribution sector at least two products can be clearly identified: number of customers, and network size. Number of customers is an output that involves customer service activities, including meter reading, billing, payment collection, and customer support. Therefore, utility companies produce "served customers". The second product is associated to maintaining and operating the network. This output requires tree pruning, combatting non-technical losses (theft), optimizing technical losses, to list just a few of the network operation and maintenance related tasks. The output in this case is identified as kilometers of operated and maintained network. It is worth noting that in addition to these products, there are some exogenous or non-managerial variables that influence costs. Candidate environmental variables include customer density and socioeconomic complexity. These environmental (or context) variables must be taken into account in the analyses in order to adjust the inputs used by the companies for conditions beyond managerial control. Finally, input prices such as wages and the cost of capital need to be included in the estimation.

Quality is a variable that could be viewed as another output dimension. Higher quality levels require managerial effort (and other resources): tasks need to be performed in conjunction with investments in order to achieve greater reliability and improved customer relations. However, under an incentive regulation environment, maximum prices are set subject to minimum quality of service standards. If a company provides a higher than required quality level, is would not be fair to account for the additional inputs that this

requires. In other words, companies providing higher than regulated quality standards are wasting resources and should not be compensated by adding the variable quality as an output. If the research asks a different question, to discover the amount of inputs that is required to produce a certain quality level, then this variable should be included. If the objective is to determine efficiency under the existing regulatory conditions, quality should not be taken into account. This study excludes quality levels in the estimated models.

3. Methodology

The methodology used in this study is a three-stage approach based on the framework applied by Fried et al (2002). This method combines the advantages of DEA and SFA methods. A complete analysis of the advantages of using a three stage model can be found at Fried et al (2002). In the first stage, we apply DEA to input and output data to obtain the initial efficiency scores of firms. In the second stage, SFA is used to attribute variation in first stage electricity distribution companies' performance to environmental variables, managerial inefficiency and random noise. In the third stage, firms' inputs are adjusted to take into account the environmental effects and statistical noise uncovered in the initial DEA. Then DEA is applied to the adjusted inputs to obtain improved measures of managerial efficiency, net of the effects of environmental variables and random noise. The procedures used in each stage is explained below:

Stage 1: The initial DEA Producer Performance Evaluation

The initial efficiency scores are obtained using a conventional DEA analysis adopting an input orientation since electricity distributors only can manage their inputs to fulfill a given demand. That is, in their concession area they are responsible for providing electricity service to every consumer connected to their distribution network. Electricity distributors are the only producers in their concession areas, since this industry is considered a natural monopoly. An industry is a natural monopoly if strict cost subadditivity is verified. Since subadditivity does not necessarily imply economies of scale, variable returns to scale was utilized in the estimation process. Formally, for the producer "o" the problem can be expressed as follows:

$$\min_{\theta,\lambda} \ \theta \qquad (1)$$

$$\begin{aligned} subject\ to\ \theta x^0 &\geq X\lambda \\ Y\lambda &\geq y^0 \\ \lambda &\geq 0 \\ e^T\lambda &= 1 \end{aligned}$$

where $x \geq 0$ is a producer´s N x 1 vector of inputs¸ y≥ 0 is a producer´s M x 1 vector of outputs, X = [x_1, . . . , x_I] is an N x I matrix of inputs vectors in the comparison set, Y = [y_1, . . . , y_I] is an M x I matrix of output vectors in the comparison set, λ = [λ_1, . . . , λ_I] is an I x 1 vector of intensity variables, e = [1, . . . , 1] is an I x 1 vector, and there are I producers in the comparison set. The data of the producer being evaluated are superscripted "o", and the problem is solved I times, once for each producer in the comparison set.

The optimal solutions to the problem formulated above are $\theta \leq 1$, which provide an initial performance evaluation for each firm, attributable to managerial inefficiencies, random noise and contextual variables.

Stage 2: Decomposing Stage 1 Slacks

In this stage, Stage 1 slacks are decomposed into environmental effects, statistical noise and managerial inefficiencies using SFA. Stage 1 total input slacks, s_{ni} are obtained as

$$s_{ni} = x_{ni} - X_n\lambda \geq 0, n = 1, \dots, N, i = 1, \dots, I$$

Where s_{ni} is the Stage 1 slack in the usage of the nth input for the ith producer, X_n is the nth row of X and $x_n\lambda$ is the optimal projection of x_{ni} onto the input efficient subset for output vector y_i.

To break down the slacks into the three effects mentioned above, a slack stochastic frontier is specified as follows:

$$s_{ni} = f^n(\mathrm{z_i}; \beta^\mathrm{n}) + \mathrm{v_{ni}} + \mathrm{u_{ni}}, \;\; \mathrm{n} = 1, \dots, \mathrm{N}, \mathrm{i} = 1, \dots, \mathrm{I}, \quad (2)$$

where $f^n(\mathrm{z_i}; \beta^\mathrm{n})$ is the deterministic feasible slack frontiers with parameter vectors β^n to be estimated, the z_i, $i = 1, \dots, K$ represent contextual exogenous variables, the v_{ni} represent the random noise, being $f^n(\mathrm{z_i}; \beta^\mathrm{n}) + v_{ni}$ the stochastic feasible slack frontier. It is assumed that $v_{ni} \sim (0, \sigma^2_{vn})$. The stochastic feasible slack frontier represents the minimum slack compatible with the variables z_i in a noisy environment. Consistent with a stochastic cost frontier, the $u_{ni} \geq 0$ and reflect managerial inefficiency. By assuming $u_{ni} \sim N^+(0, \sigma^2_{un})$ and the u_{ni} are distributed independently of each other and of the z_i, each of the N regressions of (2) can be estimated by maximum likelihood techniques and u_{ni} can be obtained by using the Jondrow et al (1982) methodology. It is be worth noting that all parameters (β^n, σ^2_{vn}, σ^2_{un}) are allowed to vary across each slack regression. This feature of the methodology is especially important when firms utilize more than one input which are affected by contextual variables in different ways; the inputs for this study are capital costs and operating cost. Moreover, with this approach it is possible for a company to be more efficient with the utilization of one input than the other.

The results of the estimations of (2) are utilized to "level the playing field" before repeating the DEA analysis in the third stage. To do so, the inputs of firms who have been advantaged by their favorable operating context and / or by their good luck are adjusted upward, according the following expression:

$$x^A_{ni} = x_{ni} + \left[max_i\{z_i\hat{\beta}^n\} - z_i\hat{\beta}^n\right] + [\max\{\hat{\mathrm{v}}_{\mathrm{ni}}\} - \hat{\mathrm{v}}_{\mathrm{ni}}], n = 1, \dots, N, i = 1, \dots, I \quad (3)$$

Where x^A_{ni} and x_{ni} are adjusted and observed inputs quantities respectively. Applying (3) to all inputs, all companies are situated in a common operating environment.

Stage 3: Adjusted DEA

In this stage, the DEA analysis is repeated, replacing x_{ni} with x^A_{ni}. Inefficiency scores from Stage 3 represent technical inefficiency attributable to managerial performance net of operating environment and statistical effects.

Data

The data used in this study are those used by the Brazilian National Electricity Agency (ANEEL) to determine the methodology that will be applied for calculating the electricity rates for the third regulatory period. This database was used by ANEEL, in Public Hearing 40, for estimating the efficiency of the distribution companies. The input and output information was provided by the companies to the regulator. The environmental variables

were obtained by ANEEL from sources described below. In addition, we segregated the total number of consumers by type in small, industrial and rural with information from various public sources.

ANEEL's database is available at its official website and it includes 61 electricity distributors for the period 2004 – 2009. Years before 2003 were not considered due to data limitations and to exclude the period of energy rationing that took place in 2001 and 2002. In this study 17 of these companies were excluded due to lack of specific data mainly related to consumer type disaggregation. Note that the firms included serve in total 59.8 million consumers, which represents 95% of the total of Brazilian consumers. Details about the definition of the variables can be found in the Technical Note 294 by ANEEL.

The variables used in this study are the following:

Operating and maintenance costs (OPEX) included the following items:

- Personnel
- Administrative costs
- Materials
- Third party services
- Rents and leases
- Insurances
- Taxes
- Others

The OPEX data come from accounting sources (Accounting Handbook of Public Service for Electricity). Costs related to services provided to third parties and those corresponding to generation and transmission activities were excluded, as well as those related to unregulated activities. Figures are expressed in local currency as of December 2010. To do this, we used the producer price index (IGPM), except for items Third Party Services and Personnel, which were adjusted with the consumer price index (IPCA) following identical criteria as ANEEL.

Capital Cost used in this study was calculated according to an economic concept. Thus, we included the following:

- Maintenance Cost of Capital: Annual Depreciation of physical capital
- Opportunity Cost of Capital: Annual Payment (dividend and interest payment) for the asset

Traditionally, the annual cost of capital can be calculated in two ways: as the annuity required for recovering the capital invested during the life of the physical assets, or as the sum of depreciation on gross physical capital plus the remuneration on net invested capital. While the annuity is a constant amount, the cost of capital obtained according to the second approach decreases over time, because of the compensation on net capital decreases. However both criteria ensure the recovery of the total capital invested in the useful life of the assets.

Using the first method, the cost of capital is independent of time, while with the second method the cost of capital will decrease over time suggesting that the company is becoming

more efficient producing the same amount of product which is misleading. Thus the second approach "punishes" companies that have more new assets because their opportunity cost of net assets is higher than the other and therefore its efficiency score will be lower.

Since the aim of this work is to measure the input-oriented efficiency, it seemed more appropriate to calculate the cost of capital as the annuity, because allocating the cost of investments evenly over the period does not harm the efficiency of companies whose net assets are newer, nor does it reward with low cost (of capital) those with older assets. For the calculation of the annual rate, we utilized an opportunity cost of capital of 16.07% and the average life of the assets of each company. Both parameters were used by the regulator in its estimates of efficiency.

Electricity delivered (in MWh): Electricity delivered is referred to the entire billed market and comes from the Tracking System Market Information for Economic Regulation of ANEEL (SAMP).

Consumers includes Residential, Commercial, Industrial, Rural and other categories in December of each year of the period.

Distribution network length (in kilometers) is based on the length of Low, Medium and High Voltage networks at the end of every year of the period.

***Average wage*:** The salary was calculated by ANEEL for its estimates of efficiency and it is expressed in local currency as of December 2010. This variable measures the labor cost a distribution company faces when hiring employees.

Complexity Index measures the degree of adversity faced by distributors in regards to electricity combatting non-technical losses. This index, prepared by ANEEL, includes of the following dimensions:

- Violence: using the number of violent deaths per 100,000 inhabitants in the area of the company´ performance as a proxy,
- Inequality: measured by the percentage of household heads in the distributor´s concession area who earn incomes at or below the minimum wage,
- Deprivation: percentage of precarious housing in the distributor´s concession area
- Infrastructure: coverage of fresh water,
- Commitment of income: measured by slowness of payment in the credit sector.

The weights of the variables included in the complexity index were obtained by regression and are those used by ANEEL[5].

Rainfall (millimeters per year) The geocoded data regarding isohyets, which composed the calculation of the index rainfall, come from the National Water Agency - ANA.

Table 1 summarizes the main characteristics of the variables defined above for 53 companies included in the analysis.

[5] See Technical Note N. 271/2010 by SRE/ANEEL.

Table 1 Descriptive statistics of the 53 electricity distributors (2004-2009)

Variable	Unit	Mean	S. D.	Minimum	Maximum	Cases	Missing
Operating cost	Million Reáis	285.5	327.5	4.5	1774.8	245	0
Capital cost	Million Reáis	503.1	648.9	1.7	3161.5	245	0
Network	Thousand Km	61.3	76.4	0.4	460.6	245	0
Total Consumers	Thousand	1462.3	1586.7	13.0	7295.2	245	0
Small Consumers	Thousand	1376.8	1471.9	12.7	6685.3	245	0
Industrial Consumers	Thousand	12.1	17.2	0.1	74.2	245	0
Rural Consumers	Thousand	73.4	97.6	0.2	535.6	245	0
Energy distributed	GW/h	76359.9	9346.8	68.1	40260.0	245	0
Average wage	Reáis/month	3285.0	631.7	1925.9	5904.7	240	5
Socioeconomic complexity index		0.18	0.11	0.02	0.46	240	5
Rainfall index	mm/year	1457.5	333.0	808.0	2450.0	240	5

Source: Own elaboration from ANEEL's data

4. Estimated models

Model 1(Total Cost): *In order to estimate the efficiency of Brazilian electricity distributors* during 2004-2009, three alternative models were estimated, which are summarized in Table 2. The first model includes as inputs, operating costs and capital costs, responding to the premise that the measurement of efficiency must take a comprehensive approach, including all inputs involved in the production process jointly. This procedure reduces errors from measuring the efficiency of operation costs conditioned on the size and quality of the distribution network, since the circumstances may differ significantly among compared companies. In other words, the approach avoids the problems associated with the substitution between two different production inputs (trade-offs). The outputs considered are the number of consumers and network length per year. For the correction of the slacks of both inputs (stage 2), we considered the following environmental variables:

- Average wage,
- Socioeconomic complexity index, and
- Rain

The expected signs of the estimated coefficients associated with the environmental variables in the adjustment of the optimal slack by SFA were as expected. Average wages are expected to have a positive impact, because the distributors that operate in areas where labor receives higher wages necessarily face higher costs than those located in areas where

the formal labor market is more depressed. The former face higher wage levels that raise their costs, a difference not attributable to inefficient management.

It is expected that companies located in areas with high proportion of social vulnerability face greater difficulties in reducing non-technical losses. It is therefore likely that operating costs to combat fraud and costs of anti-fraud network are directly related to indices of socioeconomic complexity. Accordingly, this variable is expected to be positive in the optimum adjustment of the slack.

The coefficient associated with rain is expected to be positive, since companies operating in rainy areas are exposed to higher operational costs for service restoration, including network reconnection costs, posts, and line replacement.

Model 2 (Disaggregated Customers): The second model includes the same inputs and outputs as the first, with the difference being that the consumers are broken into three categories: small (composed of residential and commercial consumers), industrial, and rural consumers.

Model 3 (OPEX Only): The third model considers as input only the operation and maintenance expenses. Unlike the previous models, it only focuses on the OPEX efficiency, by controlling the magnitude of invested capital with the network length. It is worth mentioning that the Brazilian regulatory benchmarking techniques used to promote efficiency in the last review conducted last year were applied only to OPEX. We defined outputs as the number of served customers, the total energy delivered, and the network length. The latter is also treated as an output by the Brazilian regulator, as well as by other authors, who consider the number of kilometers of maintained network as an output. Estache et al (2010), however, show that in numerous studies of benchmarking, network length is often used as input, as a proxy for capital input, usually in the estimates of cost functions or distance functions. (p. 143). Environmental variables used in the adjustment of the optimal slacks are the same as in the previous models.

Table 2 Models estimated

	DEA	SFA
Model 1 (Total Cost)	Outputs: Total Consumers, Network Length, Total Energy Delivered Inputs: OPEX, Capital Costs	Average Wage, Socioeconomic Complexity Index Rain
Model 2 (Disaggregated Customers)	Outputs: Small Consumers, Industrial Consumers, Rural Consumers, Network Length, Total Energy Delivered Inputs: OPEX, Capital Costs	Average Wage, Socioeconomic Complexity Index,
Model 3 (OPEX Only)	Outputs: Small Consumers, Industrial Consumers, Rural Consumers, Network Length, Total Energy Delivered Inputs: OPEX	Average Wage, Socioeconomic Complexity Index, Rain

Results and discussions

Estimations were performed using the LIMDEP 9.0 software. In the initial stage of the estimation the efficiency of companies in the sample was calculated adopting an input orientation DEA technique under the assumption of variable returns to scale. We chose variable returns to scale, a general assumption, because there are no a priori reasons to assume that electricity distributors face constant or decreasing average costs. In fact, the regulation of electricity distribution companies is based on that those are natural monopolies. The condition for a company to constitute a natural monopoly is the existence of strict cost subadditivity, which does not necessarily imply or require that economies of scale are verified (Sharkey, 1982 and Baumol et al, 1988). A summary of Stage 1 results is presented in Table 3.

Table 3 First stage results

Model	Outputs	Inputs	Mean	Std. Deviation	Minimum	Maximum	n
Model 1	Total Consumers, Network Length, Total Energy Delivered	OPEX, CAPEX	80.44%	0.1424	0.5001	1.000	240
Model 2	Outputs: Small Consumers, Industrial Consumers, Rural Consumers, Network Length, Total Energy Delivered	OPEX, CAPEX	85.98%	0.1336	0.5170	1.000	240
Model 3	Outputs: Small Consumers, Industrial Consumers, Rural Consumers, Network Length, Total Energy Delivered	OPEX	74.05%	0.1884	0.2800	1.000	240

From the results of the first stage, we calculated each company's slacks, which were adjusted to incorporate the effect of environmental and stochastic variables. A regression model was run for each of the slacks for each input of the model under the assumption that the error term that captures inefficiency presents a normal distribution truncated at 0. We adopted a panel data model with random effects, in which inefficiency varies in the period according to the Batese-Coelli model. The results of this stage are summarized in Table 4. As shown, in all the cases the estimated coefficients have the expected sign and are statistically significant. The variable rain only was statistically significant in Model 3, implying that excessive rainfall mainly affects operating costs. Companies operating in areas where the wage is higher face higher costs, as well as those that operate in higher social conflict areas. The parameter lambda is statistically different from zero, indicating that managerial inefficiency explains the variability of slacks between sample firms and time.

Table 4 Stochastic frontier estimation results (standard errors in parenthesis)

	Model 1		**Model 2**		**Model 3**
Independent Variable	OPEX slack	Capital Costs slack	OPEX slack	Capital Costs slack	OPEX slack
Constant	-.140D+09* (.231D+08)	-.265D+09* (.061D+09)	-.132D+09* (.250D+08)	-.187D+09* (.470D+08)	-.225D+09* (633D+08)
Average wage	31412*	.054*D+09	27095*	43091*	16882*

	(4889)	(.012D+09)	(3858)	(9959)	(7176)
Socioeconomic complexity index	.143D+09** (.689D+08)	.279D+09 *** (.163D+09)	.165D+09* (.669D+08)		.227D+09** (.136D+09)
Rain					65413* (32183)
σ_u^2	.654D+08 (.636D+23)	.141D+09 (.734D+24)	.623D+08 (.510D+23)	.155D+09 .871D+24	.122D+09* (.228D+08)
λ	2.58* (.068)	3.12* (.072)	3.12* (.059)	4.47* (.05)	4.07* (1.63)
Eta	.079* (.014)	.0816* (.0144)	.0777* (.0157)	.074* (.015)	
Log-likelihood function	-4482.6	-4626.5	-4432	-4575	-4538
N	240	240	240	240	240

*Significant at the 1% level or better

**Significant at the 5% level or better

***Significant at the 10% level or better

The impacts of environmental variables and the stochastic variables estimated in the second stage were incorporated into the inputs of each company, which were adjusted so that all distribution utilities faced a common operating environment. To do so, input quantities of the producers who were benefited from a favorable environment and "good luck" were adjusted upward according to (3).

Finally, the efficiency scores were recalculated using the inputs adjusted based on the second stage results. A summary of the results is presented in Table 5, which shows the average efficiency scores for the total sample and for some groups of distributors. As expected, in all cases efficiency is higher after separating the non-managerial component of total slacks. In general, the average efficiency rises from 0.8/0.86 to 0.93/0.92 in the estimation considering OPEX and Capital Costs as inputs (Models 1 and 2), whereas in model 3, the initial mean score, 0.74 increases to 0.89 in the last stage.

Table 5 Average Efficiency scores (Stage 3)

		Model 1		**Model 2**		**Model 3**	
Firm group	N	Stage 1	Stage 3	Stage 1	Stage 3	Stage 1	Stage 3
Smallest size Less than 500000 consumers	78	0.76	0.92	0.87	0.92	0.73	0.87
Medium Size More than 500000 and less than 3 million consumers	126	0.80	0.92	0.83	0.92	0.70	0.88
Largest size More than 3 million consumers	36	0.91	0.97	0.93	0.97	0.92	0.99
All the simple	240	0.80	0.93	0.86	0.92	0.74	0.89

Note that the average score of the companies serving larger quantities of customers (more than 3 million customers) is higher than those classified as small and medium companies both in the first and third stage of the estimations in all models, while no significant disparities are observed between the smallest and medium firms. Notice that after "leveling the playing field," according to Fried et al (2002), the distribution company' size seems to reduce its impact in the final estimations of efficiency. However, results suggest that larger

companies present advantages compared to smaller ones, attributable to higher density, reducing the unit cost.

It is also worth noting that Model 3 shows lower efficiency scores than Models 1 and 2, suggesting that when the analysis is more complete, considering both OPEX and Capital Costs, the production tradeoffs between capital and labor are better captured by the model. On the other hand, when only estimating OPEX efficiency, companies that are more labor intensive might be determined to be inefficient. It seems clear that when using both inputs, inefficiency is better isolated from the other variables that affect costs.

It would be interesting to compare the results of this study with those obtained by other authors. While there are several efficiency studies applied to, making comparisons is difficult. First, there is the time period under analysis where the majority was made with data covering earlier periods, except that made by ANEEL in the third tariff review. Another limitation on comparisons is related to the approach taken by the authors. This study estimated the efficiency using inputs in monetary terms, because of cost efficiency is the characteristic that is considered in tariff reviews that determine the utility revenue requirements. However, it is common to find in the literature studies that approximate the inputs with the number of employees and the length of the network, as Ramos Real et al (2009) and Tovar et al (2011). This approach has the disadvantage that it does not incorporate the effect of the price of raw material, especially as regards to network extension. Finally, not all studies considering costs as inputs include all kind of costs, but only OPEX. For the above reasons, it is possible to relate our results only with those obtained by ANEEL with 2003-2009 data: results applied to the current tariff cycle. There are some similarities between the results of the model by ANEEL and our Model 3, which includes only the OPEX as input. Eleven distribution companies out of the most efficient according to three-stage DEA were also ranked among the most efficient by ANEEL[6], these companies are AES-SUL, CEMAR, COELBA, COELCE, PIRATININGA, CPFL PAULISTA, ELETROPAULO, RGE, CELTINS, LIGHT and CELPE.

When comparing the results considering worst performers we also obtain similar results. Eleven distribution companies coincide with ANEEL's assessment[7].

Even though we find some similar results when identifying best and worst performers, our efficiency estimates are considerably higher than ANEEL's. In fact we obtained an average score of 90% using Model 3, while ANEEL's average is 61%. So if firms are "punished" for being far from the efficiency frontier, the ANEEL results could unfairly penalize relatively strong performers. This difference could be due to the inclusion of random effects in the frontier estimates and the breakdown of the output variables that add precision to our analysis.

5. Conclusions

In this study we estimate the efficiency of electricity distribution utilities in Brazil for 2004 - 2009, combining DEA methodology and SFA. The results show that the impact variables

[6] All companies that obtained a score above average were considered among the most efficient. For year 2009, ANEEL's estimation shows 24 out of 59 companies above this threshold while ours shows 16 out of 34.
[7] All companies that obtained a score below average were considered among the least efficient. For year 2009, ANEEL's estimation shows 35 out of 59 companies above this threshold while ours shows 18 out of 34.

beyond managerial control, such as level of input prices and socioeconomic complexity, as well as statistical noise, significantly affect the measured efficiency of firms in the sample. This effect is more significant among small and medium firms. A contribution is the incorporation of the stochastic variables in the models, in contrast with the methodology used recently by the Brazilian regulator, which explicitly included environmental variables but excluded the effect of statistical noise. Another interesting aspect is the inclusion of capital costs in the models jointly with OPEX, capturing the effect of possible trade-offs between these two inputs. Moreover, the efficiency scores of each electricity distributor show small variations over the time period, unlike the results obtained by the regulator in the last price revision, which showed sharp fluctuations in the scores for some utilities.

In addition, this study considered a more complete model of customers (Model 2). The disaggregation of consumers takes into account the different degrees of effort required to serve different user types, a feature not taken account in most efficiency studies of the electricity distribution sector.

The results indicate a trade–off between OPEX and CAPEX, since the average score increases when the inputs representative of capital are explicitly incorporated into the model. It should be noted that the frontier was recalculated after correcting the inputs (recognizing the impacts of environmental variables); other methodologies only correct the scores. This process is known as re-evaluation of the frontier. We recognize that it is necessary to explore other models and specifications in order to reflect with greater precision the production process, i.e. to consider energy delivered by voltage level or the maximum demand. As with most empirical studies, the results open up new avenues for research.

Ultimately, the issue is the sensible use of efficiency scores and rankings. When multiple methodologies yield broadly similar results, one can be more confident that the regulatory decisions are based on reality (Berg and Lin, 2008). This study is offered as a foundation for further work. In addition, analyzing results over a number of years is important, since a utility could reduce maintenance expenditures to increase its efficiency score associated with OPEX. Such behavior has negative consequence for future years, so consistency over time (such as obtained in this study) gives some confidence in the results. Without robust results, regulators must be careful when there is real money on the table. For example, the models presented here did not include service quality as an output since it was assumed that utilities were meeting minimum quality standards. However, if quality improvements need to be incentivized this important dimension of output would need to be added to the analysis (Lin and Berg, 2008). Finally, determining catch-up times and stretch factors is more an art than a science at present. If calculated distance from the frontier is used to establish targets or to penalize firms for poor performance, then the model that is used can make a huge difference in regulatory rate proceedings.

References

Banker, R.D. and R. Natarajan (2008). Evaluating contextual variables affecting productivity using data envelopment analysis. Operations Research 56(1): 48-58.

Baumol, W., Panzar, J. and Willig, R. (1988) Contestable markets and the theory of industry structure, Harcourt Brace Jovanovich, Inc.

Berg, Sanford and Chen Lin (2008). “Consistency in Performance Rankings: The Peru Water Sector,” *Applied Economics*, 40 (6), pages 793-805.

Burns, P. and Weyman-Jones, T.G, (1996). "Cost Functions and Cost Efficiency in Electricity Distribution: A Stochastic Frontier Approach," Bulletin of Economic Research, Wiley Blackwell, vol. 48(1), pages 41-64, January.

Farsi, M. and Filippini, M. (2004). ‘Regulation and Measuring Cost Efficiency with Panel Data Models: Application to Electricity Distribution Utilities’, Review of Industrial Organization, 25 (1): 1-19.

Farsi, M., Filippini, M. and Greene,W. (2005). "Application of Panel Data Models in Benchmarking Analysis of the Electrivity Distribution Sector," CEPE Working paper series 05-39, CEPE Center for Energy Policy and Economics, ETH Zurich.

Fried, H.O.; Lovell C.A.K.; Schnidt S.S.; Yaisawarng S. (2002), Accounting for Environmental Effects and Statistical Noise in Data Envelopment Analysis, Journal of Productivity Analysis 17, 157-174.

Goto, M. and Tsutsui, M. (2008). "Technical Efficiency and Impacts of Deregulation: An Analysis of Three Functions in U.S. Electric Power Utilities During the Period From 1992 Through 2000," Energy Economics, 30:1, 15-38.

Growitsch, C., Jamasb, T. and Pollitt, M. (2005). "Quality of Service, Efficiency, and Scale in Network Industries: An Analysis of European Electricity Distribution," IWH Discussion Papers 3, Halle Institute for Economic Research.

Hess, B. and Cullmann, A. (2007). "Efficiency analysis of East and West German electricity distribution companies - Do the "Ossis" really beat the "Wessis"?," Utilities Policy, Elsevier, vol. 15(3), pages 206-214, September.

Jamasb, T. and Pollitt, M., (2003). "International benchmarking and regulation: an application to European electricity distribution utilities," Energy Policy, Elsevier, vol. 31(15), pages 1609-1622, December.

Lin, Chen and Sanford Berg (2008). “Incorporating Service Quality into Yardstick Regulation: An Application to the Peru Water Sector,” *Review of Industrial Organization*, 32, (1), 2008, pages 53-75.

Mota, R., (2004). Comparing Brazil and USA electricity distribution performance: what was the impact of Privatization? Cambridge Working Papers in economics, CWPE 0423. The Cambridge-MIT Institute.

Neuberg, L. (1977) “Two issues in the municipal ownership of the electric power distribution systems”, The Bell Journal of Economic Vol. 8, No. 1 (Spring, 1977), pp. 303-323.

Pérez-Reyes, R., and Tovar, B. (2009): Measuring Efficiency and Productivity Change (PTF) in The Peruvian Electricity Distribution Companies after Reforms, Energy Policy, 37, 2249-2261.

Ramos-Real, F., Tovar, B., Iootty, M, Fagundes de Almeida, E. and Queiroz, H. (2009). "The evolution and main determinants of productivity in Brazilian electricity distribution 1998-2005: an empirical analysis", Energy Economics, 31 (2) 298-305.

Sharkey, W. (1982) The theory of natural monopoly, Cambridge University Press.

Silva, Hamilton (2011). "Cost Efficiency in Periodic Tariff Reviews: The Reference Utility Approach and the Role of Interest Groups." University of Florida, Department of Economics, PURC Working Paper.

Simar, L. and Wilson, P. W. (2007). "Estimation and inference in two-stage, semi-parametric models of production processes," Journal of Econometrics, Elsevier, vol. 136(1), pages 31-64, January.

Tovar, B., Ramos-Real, F. and Fagundes de Almeida, E. (2011). "Firm size and productivity. Evidence from the electricity distribution industry in Brazil", Energy Policy, 39 826-833.