Socio-Economic Planning Sciences 49 (2015) 47–60

Contents lists available at ScienceDirect

Socio-Economic Planning Sciences

journal homepage: www.elsevier.com/locate/seps

Statistical evaluation of Data Envelopment Analysis versus COLS Cobb–Douglas benchmarking models for the 2011 Brazilian tariff revision

Marcelo Azevedo Costa [a,*], Ana Lúcia Miranda Lopes [b], Giordano Bruno Braz de Pinho Matos [c]

[a] Department of Industrial Engineering, Universidade Federal de Minas Gerais, Belo Horizonte, MG 31270-901, Brazil
[b] CEPEAD, Universidade Federal de Minas Gerais, Belo Horizonte, MG 31270-901, Brazil
[c] CEMIG, Belo Horizonte, MG, Brazil

ARTICLE INFO

Article history:
Available online 25 November 2014

Keywords:
Benchmarking
Corrected Ordinary Least Squares
Cobb–Douglas production function
Data Envelopment Analysis

ABSTRACT

In 2011, the Brazilian Electricity Regulator (ANEEL) implemented a benchmarking model to evaluate the operational efficiency of power distribution utilities. The model is based on two benchmarking methods: Data Envelopment Analysis (DEA) and Corrected Ordinary Least Squares (COLS) with a Cobb Douglas production function. Although the estimated scores are highly correlated, differences between the scores are as high as 41%. For some companies differences between the efficiency scores result in substantial reduction in regulatory operational costs. We provide a detailed statistical comparison which indicates that the COLS Cobb Douglas model has major deficiencies in terms of estimating efficiency scores.

1. Introduction

The use of frontier-based methods in the incentive regulation of power companies has grown in recent years. With an international survey performed in 40 countries on a total of 43 regulators between June and October 2008, Haney and Pollitt [1] found that 51% of the companies were applying benchmarking techniques. The frontier techniques can be divided into three major groups: Data Envelopment Analysis (DEA), Corrected Least Squares (COLS) analysis and Stochastic Frontier Analysis (SFA). The authors found that in 2008, 34.8%, 13% and 8.7% of the power distribution regulation models were applying DEA, COLS analysis and SFA, respectively. They also found that 40%, 15% and 5% were applying DEA, COLS and SFA, respectively, in power transmission regulation models. Haney and Pollitt [1] report that Austria applies DEA and COLS for both power transmission and distribution utilities; Belgium applies DEA for power transmission and distribution and applies SFA and COLS in power transmission regulation models; Denmark applies COLS analysis; Portugal applies SFA; Slovenia applies DEA; Iceland, Netherlands and Norway apply DEA in power distribution and transmission models; Argentina and Brazil applied DEA for power transmission in 2008, but Brazil applied DEA and COLS models for power distribution companies in 2011. Colômbia began to apply DEA for power transmission in 2000 and for power distribution in 2002.

Several studies have indicated the suitability of DEA in the analysis of efficiency in power-regulated sectors. For example, Edvardsen et al. [2] describe the Norwegian Electricity regulation model. The Norwegian Water Resources and Energy Directorate (NVE) was one of the first European regulators to use DEA. Bogetoft and Otto [3] describe the DEA-based incentive model applied in Germany; Hu and Wang [4] evaluate the efficiency of electric utilities and their effects on consumer prices; Souza et al. [5] compare DEA and SFA in the measurement of the efficiency of 40 Brazilian energy distribution companies; Sarıca and Or [6] assess the efficiency of Turkish power plants using DEA; Vaninsky [7] uses DEA to measure the efficiency of electric power generation in the United States; Hu and Wang [4] analyze the energy efficiencies of 19 administrative regions in China for the period of 1995–2002 using DEA. A summary of the use of DEA in energy and environmental studies can be found in Ref. [8].

In September 10, 2010, the Brazilian National Electric Energy Agency began a debate with the Brazilian society regarding the rules and methodologies for defining the revenues of electricity distribution utilities for the 3rd Periodic Tariff Review Cycle (3PTRC) through public hearing 040/2010 (AP040). Through

* Corresponding author.
E-mail address: macosta@ufmg.br (M.A. Costa).

http://dx.doi.org/10.1016/j.seps.2014.11.001
0038-0121/

technical note 265/2010, the national regulator proposed a full review of the model that calculates regulatory operational costs. The definition of efficient operating cost is a central point in incentive regulation because it was chosen for the regulation of natural monopolies in the Brazilian electricity sector after their privatization in the nineteen-nineties. Incentive regulation requires the definition of a revenue or rate level for a fixed period of time, which is defined in a formal contract. Given the rate or revenue level defined by the regulator, companies are encouraged to reduce their costs to achieve higher financial returns. After the concession period defined in the agreement contract, which is 4 or 5 years, the companies' costs are revised; i.e., the regulator defines new levels that are considered more efficient. In this case, the new efficiency levels are proposed by the regulator for the benefit of consumers. Therefore, a consistent methodology for the definition of operational costs is important for both the regulated company and consumers.

This paper evaluates the benchmarking models proposed by ANEEL, which are the DEA and COLS models, and discusses major inconsistencies of the COLS methodology. Specifically, we aim to identify the causes of the discrepancy between the DEA and COLS methodologies. We provide a detailed statistical analysis of the Cobb Douglas regression model and the efficiency score estimates. We apply statistical hypothesis testing, linear regression theory, and Monte Carlo simulations. The results show that the differences between the COLS Cobb–Douglas and the DEA efficiency scores are statistically significant. We show that the main causes of the differences are related to sample size. We also explore missing variables in the Cobb–Douglas regression model.

The paper is organized as follows: Section 2 presents the historical background of the benchmarking models proposed by ANEEL in 2011, reviews the DEA model and the Cobb–Douglas production function, and presents the statistical methods used to evaluate the efficiency scores. Section 3 presents the results. Section 4 discusses the main findings. Section 5 presents the conclusion.

2. Material and methods

2.1. Background

During an initial stage of public hearing 040/2010, ANEEL proposed two DEA models organized into two stages, hereafter named model 1 and model 2. In the first stage, both models use operational cost as the input variable and network extension (km – kilometers) as the output variable. Model 1 uses number of customers as the second output variable, whereas model 2 uses energy consumption (MWh – Mega Watt-hour) as the second output variable. Both models assume non-decreasing returns to scale (NDRS).

The Data Envelopment Analysis (DEA) benchmarking method was first considered by ANEEL because it has been successfully applied by other regulatory agencies from Austria, Britain, Belgium, Finland, Netherlands, among others. Therefore, it is of interest of ANEEL to replace previous reference company method by a concise and robust benchmarking approach, like DEA. By applying DEA, the operating cost of each company is compared to the costs/results of the remaining companies and, by means of a linear system of equations, an efficiency frontier is calculated. The final results of this model are efficiency scores that indicate the efficiency of each company in transforming inputs (costs) in outputs (electricity consumption, number of customers and network extension), when compared with similar companies.

The parameters of the models were estimated using a database covering the years 2003–2009. However, first, the power distribution companies were split into two groups. Group A is composed of companies with a 2003 annual energy consumption greater than 1 Tera Watt-hour (TWh), whereas group B is composed of companies with a 2003 annual energy consumption smaller than 1 TWh. The DEA method was applied separately to each group to estimate the efficiency scores. Later, the estimated scores were further adjusted using a second-stage model with environmental variables.

After public contributions, the benchmarking methodology was reviewed by ANEEL. The new proposed methodology consists of two benchmarking models: one DEA model in the first stage that aggregates all previous output variables and a second benchmarking model known as Corrected Ordinary Least Squares (COLS), presented in Technical Note 101/2011. COLS is a parametric model that fits a linear regression model using ordinary least squares. In sequence, the regression model is shifted toward the smallest observed value among the residuals of the regression model. Thus, the lower bound or the efficiency frontier of the operational cost is determined. In this case, the regression model equation is the Cobb–Douglas production function [9].

The COLS method is currently used by two regulators (Denmark and Great Britain), and it is known to be more restrictive; i.e., it strongly penalizes companies that are not on the frontier [3]. To overcome this limitation, ANEEL applies the mean value of the DEA and COLS efficiency score as the first stage outcome.

In ANEEL's final decision, released in November 2011, the energy consumption (MWh) output variable was replaced by a weighted energy consumption variable that aggregates high-, medium- and lower-voltage energy consumption. The weights were chosen to be proportional to the amount of consumption in the high-, medium-, and lower-voltage markets of each company. Again, the DEA and COLS methods were applied separately to groups A and B to estimate the final efficiency scores. A summary of the proposed models is presented in Table 1.

Although the methodology proposed by ANEEL was based on the experience of leading European regulatory agencies, it was subject to criticisms and suggestions from Brazilian community and power distribution companies. One of the major concerns was the use of the DEA model with non-decreasing returns to scale (NDRS) as a replacement for the most commonly used model, the DEA with variable returns to scale (VRS) [10]. A technical report was submitted to ANEEL [11] providing evidence that the VRS model is the most appropriate model. Banker [11] states that "*even if economic theory argues that non-decreasing returns to scale prevails in situations of natural monopoly, empirical evidence strongly suggests that*

Table 1
Proposed benchmarking models by ANEEL.

Method	Input variables	Output variables
Technical note 265/2010 (first proposal)		
DEA-NDRS	Operational Cost (R$)	Network length (km), Number of customers
DEA-NDRS	Operational Cost (R$)	Network length (km), Power consumption (MWh)
Technical note 101/2011 (second proposal)		
DEA-NDRS	Operational Cost (R$)	Network length (km), Number of customers, Power consumption (MWh)
COLS	Operational Cost (R$)	Network length (km), Number of customers, Power consumption (MWh)
Technical note 294/2011 (final proposal)		
DEA-NDRS	Operational Cost (R$)	Network length (km), Number of customers, Weighted power consumption (MWh)
COLS	Operational Cost (R$)	Network length (km), Number of customers, Weighted power consumption (MWh)