1-1-1999

A Comparison of Efficiency Measurement Using Data Envelopment Analysis and Regression in the Utility Industry

Janice Forrester
Cytera Systems, Inc.

Timothy R. Anderson
Portland State University, tim.anderson@pdx.edu

Citation Details

Forrester, J.; Anderson, T.R., "A comparison of efficiency measurement using data envelopment analysis and regression in the utility industry," Management of Engineering and Technology, 1999. Technology and Innovation Management. PICMET '99. Portland International Conference on, vol.1, pp.297.

A Comparison of Efficiency Measurement Using Data Envelopment Analysis and Regression in the Utility Industry

Janice Forrester[1], Timothy R. Anderson[2]
[1]Director of Strategic and Analytical Services, Cytera Systems, Inc.
9220 SW Barbur Blvd., Suite 119-148, Portland, OR 97213, USA
[2]Assistant Professor, Engineering Management Program, Portland State University
PO Box 751, Portland, OR 97207 – 0751, USA

Economic theory of productive efficiency is based on the comparative analysis of the best-in-class producers vis-à-vis all others. The criterion for determining the "best" producers refers to the ability to produce maximum output given a specific level of input, or conversely, the ability to use the least amount of input to produce a specific level of output.

Traditionally, research on efficiency has relied on one of two approaches: a parametric approach using econometric tools or a nonparametric approach using linear programming techniques, such as DEA. Econometric methods involve estimating a production function based, on average, on how various inputs are used by a group of similar producers. These techniques also require that certain statistical assumptions be satisfied (e.g., that there should exist no significant relationship among various independent variables or inputs) and some a priori knowledge of the functional form. On the other hand, DEA, being nonparametric, requires no such assumptions. DEA also optimizes each company individually (by benchmarking it against its closest peers), whereas traditional statistical methodologies rely on averages or single optimization approaches.

It is the intent of this paper to first discuss the theoretical differences between the manner in which regression and DEA separate variation into that which is caused by noise and that caused by inefficiency. Secondly, to examine the empirical results of independent studies, one that used regression and one that used DEA. The studies were performed one year apart by independent teams using the same data and responding to the same questions.

Both studies focus on monitoring the efficiency of the utility providers in light of deregulation. In the past the utility industry was secure in its quasi-monopolistic market and cost-based rate setting. Inefficiency was often rewarded and there was little incentive to improve. However, with the announcement of the deregulation of the utility industry, the market will no longer be captive, open access will allow any utility to serve any customer anywhere. At the beginning, cherry picking will be the most dangerous form of competition among the utilities. Those who are efficient and can deliver the commodity at a cheap price will win.

Both studies were performed using data obtained from POWERdat ©1998 Version 2.01, a Resource Data International, Inc., database. Original data sources included The Federal Energy Regulatory Commission (FERC) Form 1 and the US Securities & Exchange Commission 10 K and 10 Q reports for Holding Companies and Utility operations. The data set included 140 holding companies the period from 1990 to 1995 for the regression study and 1990 to 1996 for the DEA study.

The first study was a national study performed in 1997 by Haeri, Khawaja and Perussi that examined the ability of each utility company to transform resource materials into power using regression techniques. Specifically, they constructed a Cobb-Douglas production function with the dependant variable power being measured in Megawatt Hours (MWh) and the independent variables, the resources, were capital, labor, fuel, and materials. In addition, the authors included a load factor variable to account for idle capacity and a trend variable to account for changed in technology. The model was built with full access to all of the previously mentioned data from POWERdat and with sole objective of providing the best possible measure of inefficiency for each of the utilities.

In 1998 a similar national level study by Forrester, Khawaja, Haeri, and Carter was performed on the same utilities over the same time period using static and dynamic DEA techniques. The output for the model was power measured in MWh with labor, operation and maintenance, fuel, pension and benefits, and capital as the inputs. The model built with the sole intent of measuring the inefficiencies in the utility industry.

Historically, comparisons between regression models and DEA have been performed with simulated data and identical variable selection. This work represents an interesting opportunity to be able to compare these approaches based on individually published models which were given access to the same data and designed to meet the same objective, yet built independently over a year apart. The models were not built for comparative purposes, but rather to be the best possible model to evaluate the efficiency of the utility industry.