Assessing the Efficiency and Optimal Scale of Electricity Distribution Businesses in New Zealand

Introduction

This draft focuses on empirical and econometric methods relevant to benchmarking the efficiency of Electricity Distribution Businesses (EDBs). It includes a preliminary discussion on assessing optimal scale. The aim is to complement a broader report that offers a more strategic overview of the key issues.

1 Benchmarking Efficiency

For robust efficiency benchmarking of New Zealand's Electricity Distribution Businesses (EDBs), I recommend a multi-method approach, emphasising the advantages of using Data Envelopment Analysis (DEA), Stochastic Frontier Analysis (SFA), and mixed methods.

1.1 Data Envelopment Analysis (DEA)

Data Envelopment Analysis (DEA) is a non-parametric method used to evaluate the relative efficiency of decision-making units (DMUs), such as Electricity Distribution Businesses (EDBs), based on multiple inputs and outputs. DEA constructs an efficiency frontier and compares each DMU against this frontier.

DEA Model Formulation

Assume there are n EDBs, each using m inputs to produce s outputs. Let:

- x_{ij}: amount of input i used by EDB j
- y_{rj}: amount of output r produced by EDB j
- x_0, y_0: input and output vectors of the EDB being evaluated

The input-oriented DEA model under Variable Returns to Scale (VRS) is:

$$
\begin{aligned}
\min_{\theta,\lambda} \quad & \theta \\
\text{s.t.} \quad & \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{r0}, \quad r = 1, \ldots, s \\
& \sum_{j=1}^{n} \lambda_j x_{ij} \leq \theta x_{i0}, \quad i = 1, \ldots, m \\
& \sum_{j=1}^{n} \lambda_j = 1 \\
& \lambda_j \geq 0, \quad j = 1, \ldots, n
\end{aligned}
$$

The efficiency score θ satisfies $0 < \theta \leq 1$. A score of 1 indicates full efficiency.

The measure obtained from solving the linear problem above is known as the Farrell input measure of technical efficiency. This is by far the most popular measure in the efficiency benchmarking studies in the electricity distribution. There are a range of other formulations of DEA in the literature (for an overview, see, e.g., ch. 8–10 in [10])

Inputs and Outputs for EDBs

Typical inputs: Operating expenditure (OPEX), Capital expenditure (CAPEX), Network length, Transformer capacity
Typical outputs: Energy delivered (GWh, Number of customers, Reliability indices (e.g., SAIDI, SAIFI)

Attractive Features of DEA

It is by construction **nonparametric**; DEA does not assume a specific parametric functional form for the relationship between input and output.
One may even argue that it is **model free**; no (economic) model is assumed.
It is **most widely used**. (see [5], for example)

Limitations of Conventional DEA

DEA is a deterministic, nonparametric method that treats all deviations from the efficiency frontier as inefficiency. It **cannot distinguish between inefficiency and statistical noise** (measurement errors, random shocks, or external factors).
A follow up result is that it is **highly sensitive to outliers and measurement errors**.
There is no statistical framework and then **no hypothesis testing or confidence intervals around efficiency scores**.

Adressing limitations

To address this limitation, several methods have been proposed. One prominent approach is bootstrapping, which involves two steps: first, computing efficiency scores using the original DEA model; and second, estimating the standard errors of these scores through resampling

techniques [11]. While bootstrapping helps assess sampling variability, it does not fully account for random noise in the data. As noted by [4], these techniques are primarily designed to evaluate how DEA estimates might vary across different samples, rather than to model stochastic disturbances explicitly. For this reason, stochastic frontier analysis (SFA) is often preferred when noise modeling is essential.
There exists, for example, an application to assess the performance of the electricity distribution system in Australia in [8].

1.2 Stochastic Frontier Analysis (SFA)

SFA separates inefficiency from statistical noise using econometric modeling.

SFA Model

$$\ln y_i = f(x_i; \beta) + v_i - u_i$$

where $v_i \sim N(0, \sigma_v^2)$ is noise and $u_i \sim |N(0, \sigma_u^2)|$ is inefficiency.

Typical Application Steps

1. Specify a functional form (e.g., Cobb-Douglas).
2. Estimate parameters using maximum likelihood.
3. Derive firm-specific efficiency scores.

Attractive Features of SFA

SFA explicitly models a composite error term **separating inefficiency from random noise**, making it **more robust to data imperfections** common in electricity distribution data.
SFA, allows for **statistical inference**.
It imposes a functional form and parameter restrictions, which can provide **more economically interpretable efficiency estimates**. It makes it easier to incorporate environmental variables directly and **model inefficiency determinants**.

Limitations of Conventional SFA

SFA requires **specifying a functional form** (e.g., Cobb-Douglas, Translog) and **distributional assumptions** for inefficiency and noise, which if misspecified, can bias results.

Adressing limitations of Conventional SFA

Nonparametric SFA combines the flexibility of DEA with the noise-handling capability of SFA. It avoids strong assumptions about the functional form and error distributions.
StoNED (Stochastic Nonparametric Envelopment of Data) uses Convex Nonparametric Least Squares (CNLS) to estimate the frontier. Then applies a stochastic decomposition to

separate inefficiency from noise. It has been succesfully applied in Finland, see [7].
Many **recent alternatives** exist, for example a robust tractable approach can be found in [3].

1.3 Corrected Ordinary Least Squares

Historically, a popular alternative is Corrected Ordinary Least Squares (COLS). It is simple to implement using standard regression tools. Even though some recent improvements exist (see [9]), it is limited by having the worst of DEA (no separation of noise and efficiency) and conventional SFA (parametric) It can be still useful as a benchmark.

1.4 Metafrontier

It is worth noting that the metafrontier framework can be used in efficiency analysis to compare the performance of EDBs, possibly regrouped into groups, that operate under different technologies or environments. (geographical characteristics, specific transmission constraints, ...). It extends traditional frontier models like DEA or SFA by introducing a common benchmark (the metafrontier) that envelops all group-specific frontiers. An example of application in Taiwan can be found in [6].

1.5 Ease of Implementation

Conventional methods such as DEA and SFA benefit from readily available off-the-shelf software packages. See, for example, in [2].
COLS (Corrected Ordinary Least Squares) can be implemented using any standard statistical software.
In contrast, the adoption of StoNED may be limited due to its likely dependence on proprietary optimization solvers.
Other nonparametric SFA approaches, while potentially more time-consuming to implement, are generally compatible with standard open-source platforms like R and do not typically require proprietary tools.

1.6 Recommendation

I recommend implementing different methods and comparing results, including at least different DEA and SFA approaches.
As presented in [1] , results demonstrate that combination approaches, such as taking the maximum or the mean over DEA and SFA efficiency scores, have certain practical merits and might offer a useful alternative to strict reliance on a singular method. In particular, the results highlight that taking the maximum not only minimizes the risk of underestimation, but can also improve the precision of efficiency estimation.

2 Optimal Scale

I will briefly start a discussion on this topic, but I will not have time to elaborate.

2.1 A Word of Caution

Concerning the study of optimal scale and market structure for EDBs, empirical methods are less developed and robust than in efficiency studies.

Efficiency studies using DEA or SFA are not meant to create an absolute efficiency rank. The original motivation is to create a relative efficiency score, for example compared to an optimal frontier. Efficient units can be studied to identify best practices or technological advantages. Less efficient units can be guided on how to improve by comparing them to similar, more efficient peers.

2.2 Cost Function Estimation

I will briefly discuss a possible tractable method to study optimal scale for EDBs.

Purpose

Estimates how costs vary with output and environmental factors to assess scale efficiency.

Model

Translog cost function:

$$\ln C = \alpha_0 + \sum_i \alpha_i \ln y_i + \frac{1}{2} \sum_i \sum_j \beta_{ij} \ln y_i \ln y_j + \sum_k \gamma_k z_k + \epsilon$$

Scale Elasticity

$$SE = \left(\sum_i \frac{\partial \ln C}{\partial \ln y_i} \right)^{-1}$$

where z can include network or region characteristics.

Use in Regulation

It can support decisions on mergers and optimal firm size.
SE > 1: Diseconomies of scale suggests the firm may be too large.
SE < 1: Economies of scale suggests the firm may be too small.
SE $= 1$: Constant returns to scale suggests optimal size may have been reached.

Caveats

They are some standard caveats found in the literature. First, the accuracy of scale elasticity depends on the correct specification of the cost function (e.g., translog, Cobb-Douglas). Second, we actually face a multi-output problem: EDBs often provide multiple services (e.g., residential, industrial supply), so multi-output cost functions may be needed. Third,optimal size may change over time due to technology, demand shifts, or policy changes.

References

[1] Mark A. Andor, Christopher Parmeter, and Stephan Sommer. "Combining uncertainty with uncertainty to get certainty? Efficiency analysis for regulation purposes". In: *European Journal of Operational Research* 274.1 (2019), pp. 240–252. DOI: `10.1016/j.ejor.2018.10.007`.

[2] Peter Bogetoft and Lars Otto. *Benchmarking with DEA, SFA, and R*. New York: Springer, 2010.

[3] Samuele Centorrino and Christopher F. Parmeter. "Nonparametric estimation of stochastic frontier models with weak separability". In: *Journal of Econometrics* 238.2 (2024). DOI: `10.1016/j.jeconom.2023.105641`.

[4] Tim Coelli et al. *An Introduction to Efficiency and Productivity Analysis*. 2nd ed. New York: Springer, 2005.

[5] Aoife Brophy Haney and Michael G. Pollitt. "Efficiency analysis of energy networks: An international survey of regulators". In: *Energy Policy* 37.12 (2009), pp. 5814–5830. ISSN: 0301-4215. DOI: `https://doi.org/10.1016/j.enpol.2009.08.047`.

[6] Yi-Ju Huang, Ku-Hsieh Chen, and Chih-Hai Yang. "Cost efficiency and optimal scale of electricity distribution firms in Taiwan: An application of metafrontier analysis". In: *Energy Economics* 32.1 (2010), pp. 15–23. DOI: `10.1016/j.eneco.2009.03.005`.

[7] Timo Kuosmanen. "Stochastic semi-nonparametric frontier estimation of electricity distribution networks: Application of the StoNED method in the Finnish regulatory model". In: *Energy Economics* 34.6 (2012), pp. 2189–2199. DOI: `10.1016/j.eneco.2012.03.005`.

[8] Boon L. Lee et al. "Deregulation, efficiency and policy determination: An analysis of Australia's electricity distribution sector". In: *Energy Economics* 98 (2021). DOI: `10.1016/j.eneco.2021.105210`.

[9] Christopher F Parmeter and Shirong Zhao. "An alternative corrected ordinary least squares estimator for the stochastic frontier model". In: *Empirical economics.* 64.6 (2023-06), pp. 2831–2857. ISSN: 0377-7332.

[10] Robin C. Sickles and Valentin Zelenyuk. *Measurement of Productivity and Efficiency: Theory and Practice*. Cambridge University Press, 2019.

[11] L. Simar and P. W. Wilson. "Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models". In: *Management Science* 44.1 (1998), pp. 49–61.