Huegin's response to Draft Determination on behalf of NNSW and ActewAGL

Technical response to the application of benchmarking by the AER

Date: 16 January 2014

huegin

www.huegin.com.au
Level 10, 2 Elizabeth Plaza, North Sydney, NSW 2060
02 9409 2400

Contents

Introduction

This chapter outlines the report context, our letter of instructions and the credentials and experience of Huegin and its personnel who contributed to this report.

1

1.1 Context for this report

In November 2012 the Australian Energy Market Commission (AEMC) completed the Economic Regulation of Network Service Providers Rule Change[1] . This rule change required the Australian Energy Regulator (AER) to develop and publish a series of guidelines on its approach to regulating network service providers (NSPs), including the approach the AER will use to assess the efficiency of operating expenditure (opex) forecasts.

The AER is required to accept a DNSP's forecast opex where it is satisfied that the forecast opex for the regulatory control period reasonably reflects the criteria (the opex criteria) in clause 6.5.6(c) of the National Electricity Rules (NER), being:

- the efficient costs of achieving the opex objectives in clause 6.5.6(a) of the NER (opex objectives);
- the costs that a prudent operator would require to achieve the opex objectives; and
- a realistic expectation of the demand forecast and cost inputs required to achieve the opex objectives.

In deciding whether or not it is satisfied that the forecast opex for the regulatory control period reasonably reflects the opex criteria, the AER must have regard to certain factors specified in clause 6.5.6(e) of the NER, including, relevantly:

- the most recent annual benchmarking report that has been published under clause 6.27 of the NER and the benchmark opex that would be incurred by an efficient DNSP over the relevant regulatory control period (clause 6.5.6(e)(4)). Under clause 6.27 of the NER, the AER must prepare and publish an annual benchmarking report which should describe the relative efficiency of each DNSP in providing direct control services over a 12 month period;
- the actual and expected operating expenditure of the DNSP during any preceding regulatory control periods (clause 6.5.6(e)(5));
- the relative prices of operating and capital inputs (clause 6.5.6(e)(6));
- the substitution possibilities between opex and capital expenditure (capex) (clause 6.5.6(e)(7)); and
- any other factor the AER considers relevant and which the AER has notified the DNSP in writing, prior to the submission of its revised regulatory proposal under clause 6.10.3 is an operating expenditure factor (clause 6.5.6(e)(12).

The role of benchmarking in the determination of regulatory expenditure, as reflected in the recent draft decision for NSW and ACT, has changed over time. In the early guidance, the AER initially indicated that benchmarking would be used to assess the past performance of a DNSP to determine the degree to which this performance has been efficient relative to peers.

> ***While we examine revealed costs in the first instance, we need to test whether DNSPs responded to the incentive framework in place. For this reason, we will assess the efficiency of base year expenditures using our techniques, beginning with the economic benchmarking and category analysis, to determine if it is appropriate for us to rely on a DNSPs revealed costs. That is, whether the DNSPs past performance was efficient relative to its peers and consistent with historical trends.***
>
> - AER, Better Regulation, p8 "Draft Expenditure Forecast Assessment Guideline for Electricity Distribution", August 2013

It can be inferred from this statement that the initial intent of the AER was to use economic benchmarking as an input to deciding whether opex can be classed as either efficient or inefficient relative to peers and the DNSP's own historical cost trends.

[1] AEMC, Economic Regulation of Network Service Providers and Revenue Regulation of Gas Services, Final Position Paper, 2012

In the initial guidance in the Draft Expenditure Forecast Assessment Guideline, the AER also stated that base year opex would be set at actual expenditure should that actual expenditure reasonably reflect the opex criteria.

> ***The 'revealed cost' approach is our preferred approach to determining base opex. If actual expenditure in the base year reasonably reflects the opex criteria, we will set base opex equal to actual expenditure for those cost categories forecast using the revealed cost approach.***
>
> - AER, Better Regulation, p15 "Draft Expenditure Forecast Assessment Guideline for Electricity Distribution", August 2013

The Explanatory Statement for the final Expenditure Forecast Assessment Guideline, however, portends a shift in the stated use of benchmarking.

> ***We are likely to use economic benchmarking to (among other things):***
>
> 1. ***Measure the rate of change in, and overall efficiency of, NSPs. This will provide an indication of the efficiency of historical expenditures and the appropriateness of their use in forecasts;***
> 2. ***Develop a top down total forecast of total expenditure;***
> 3. ***Develop a top down forecast of opex taking into account:***
>   - ***the efficiency of historical opex***
>   - ***the expected rate of change for opex***
>
> - AER, Better Regulation, Explanatory Statement, p78-9, "Expenditure Forecast Assessment Guideline for Electricity Distribution", November 2013

This latest guidance from the AER raised the prospect that economic benchmarking would play a key role in the AER's assessment of DNSP expenditure and the degree to which it is considered efficient. Of particular interest is the indication that should a DNSP be deemed not to be responding to the incentive regime then benchmarking will be used as a means for the AER to determine a substitute forecast allowance.

1.2 Our instructions

Huegin has been asked to prepare this report by Networks NSW (comprising Ausgrid, Endeavour Energy and Essential Energy) on behalf of Networks NSW and ActewAGL. The subject of this report is the Australian Energy Regulator's (AER's) upcoming regulatory determination for these organisations, applicable from 1 July 2014 to 30 June 2019.

The scope of the engagement covers:

1. provide economic analysis and advice;
2. prepare a written expert report (or reports);
3. appear as an expert witness for Networks NSW or ActewAGL (if required); and
4. undertake such other work as NNSW or ActewAGL may instruct you as the Response progresses.

Specifically, we have been asked to address (**FOUR**) questions, these are:

a) Comment on whether there is consensus on one right benchmarking method and what are some of the common challenges in carrying out such analyses

b) Review the benchmarking approach used by the AER. In addressing the foregoing, the consultant is asked to address whether the results can be relied on in setting regulatory allowance for operating expenditure

c) Comment on the reasonableness of the AER's benchmarking approach. In addressing the foregoing, the consultant is asked to address whether the AER has adequately identified, and reflected in the underlying analysis, operating and environmental variables that may affect benchmarking results, and any other factors that may contribute to the potential for error.

d) In relation to b above, comment on whether the benchmarking analysis should be used, and whether the deficiencies highlighted in b) and c) above can be addressed.

Our complete instructions from Networks NSW are attached as Annex A to this report.

1.3 The qualifications and experience of report contributors

Huegin is a significant contributor to the body of knowledge for benchmarking as applied to businesses in the National Electricity Market (NEM). Huegin is also the benchmarking partner to the majority of businesses in the NEM.

The Huegin team has an appropriate mix of tertiary education and professional experience commensurate with the requirements of the task to review and critique the benchmarking analysis relied upon by the Australian Energy Regulator (AER) in the NSW and ACT draft decision. Qualifications and headline experience of those members who have contributed to this report include:

- Jamie Blair. BEng (Chemical): Jamie is a Director in our Sydney office. Jamie is the lead author of major domestic and international benchmarking studies for the electricity industry. Jamie provides regulatory support to numerous Distribution Network Service Providers (DNSPs) throughout Australia.
- Oliver Skelding. BA (Economics), MEc: Oliver is a Senior Analyst in our Sydney Office. Oliver has a Masters of Economics, specialising in Econometrics and is a major contributor to both the analysis and written articles on economic benchmarking relied upon by over 80% of the DNSPs operating in the NEM.
- Dr Ben Petschel. BSc (Mathematics) hons, PhD (Mathematical Finance): Ben is a Senior Analyst in our Sydney office. Ben has developed numerous models to provide analytical decision support within the electricity industry. These include models covering the drivers of cost and performance of wood pole populations throughout the NEM.
- Naomi Donohue. BBus (Accountancy and Computer Applications), CPA: Naomi is a Manager based in Brisbane. She worked for eight years in the regulatory and finance areas of a large DNSP, developing an in-depth understanding of cost structures and drivers within the regulatory construct.
- Darryl Walker. BEng (Aerospace) hons, MSc (Thermal Power): Darryl is a Director in our Sydney office. Darryl has worked with numerous electricity distribution companies on both engineering and operations improvement assignments. Recently Darryl authored a paper quantifying the differences in cost caused by exogenous factors affecting electricity networks in Australia.

Full profiles of team members are included at Annex B.

All contributors have read and understood the Practice Note CM7: Expert witnesses in proceedings in the Federal Court of Australia, June 2013. As lead author, Jamie Blair certifies that this report complies with Practice Note CM7. In accordance with the Guidelines, I confirm that I have made all inquiries that I believe are desirable and appropriate, and that no matters of significance that I regard as relevant have, to my knowledge, been withheld from the Court.

Signed	Date
	16 January 2015

1.4 Huegin expertise

Huegin focuses on providing analytical decision support which requires a knowledge of the way in which complex systems, such as electricity networks, work. Our team has significant experience in, and ongoing exposure to, operations improvement across many sectors including the electricity distribution sector. Given the ongoing drive for performance improvement in the electricity industry, a key focus in recent years has been understanding and modelling the drivers of performance and cost, as well as the degree to which businesses can influence these.

- Understanding and modelling the drivers of performance: The drivers of performance were first presented in the Australian DNSP benchmarking report in 2012. Since that time Huegin has continued to refine an explanatory model addressing the different drivers affecting Australian DNSPs. The effect of these eight drivers has been quantified and shown to significantly influence the results of benchmarking analysis.

- Understanding the degree to which drivers can be influenced: Huegin has developed a framework for explaining the degree to which organisations can influence the drivers of performance and cost. This framework highlights the need to understand the degree to which businesses can manage costs and performance when looking to assess relative performance and efficiency.

In addition to understanding and applying the benchmarking techniques as favoured by the AER, Huegin has focussed on the utility of benchmarking for supporting performance improvement decisions in the context of the Australian electricity industry.

The Huegin approach to benchmarking continues to evolve through the continued accumulation of this operational experience, application of specialist skills and research on the approaches and outcomes of benchmarking in other jurisdictions and industries.

This experience includes many benchmarking investigations on behalf of Australian DNSPs, notably a 2012 report of the costs and differences between many of the Australian networks and selected international networks. The purpose of that report was to provide a basic analysis of key issues in benchmarking and to share information amongst the businesses that was not available to them prior to the AER's publication of the Regulatory Information Notices (RINs) publicly. Note that the 2012 report and the data relied upon pre-dated the RINs in their current format and the data and analysis therefore differ from the current context, the draft decision and this report. This highlights the sensitivity of benchmarking outcomes to context, time and data which will be explored throughout this report.

Successful application of benchmarking techniques for the purposes of performance comparison and decision making requires fluency in specialist techniques. The techniques regularly used by Huegin include econometric analysis, statistical analysis and advanced mathematical techniques.

Despite benchmarking being relatively new in the context of revenue setting in Australia, it has been applied in various ways in a number of industries and jurisdictions. Huegin continues to critically review the approaches and outcomes of benchmarking as applied by organisations such as the Office of Gas and Electricity Markets (OFGEM) in the United Kingdom.

Ongoing knowledge is developed, applied and tested by Huegin in various ways including:

- The development of reports and submissions
- The completion of investigative analyses
- The ongoing development of the Conduit benchmarking portal
- Ongoing participation in industry forums

Based on the specialist knowledge developed, Huegin is able to comment authoritatively on the application and utility of benchmarking in the context of regulating Australian DNSPs operating in the NEM.

The benchmarking landscape

Huegin and others have, throughout the consultation period of the Better Regulation Program, pointed out the options, challenges and limitations in applying benchmarking to electricity network efficiency measurement. In this chapter we provide an overview of the most common techniques and briefly discuss some of the more consequential issues associated with benchmarking. Acknowledging and recognising challenges and limitations is important in the application of benchmarking. Misunderstanding the criticality of the issues when *relying* upon benchmarking results will render the outcomes meaningless.

2

In this chapter

The following key points are relevant to this chapter on benchmarking techniques

Huegin was asked to present our view on the most common benchmarking techniques used in electricity distribution, the challenges associated with the analysis and comment on whether consensus exists on the most appropriate method. There are many approaches to benchmarking. Regardless of the approach adopted, benchmarking requires subjective assumptions to be made about the specifications of models, types of cost production functions and treatment of data. Given the complex nature of operating and maintaining an electricity network, the lack of local competitors operating under the same regional influences and the limited amount of data available, electricity network benchmarking remains more art than science.

Key Point

There are many forms of benchmarking. Most have their origins in other industries, where outputs are more tangible and comparable across competitive markets.

In the electricity industry there is no consensus on the most appropriate form of modelling or technique. Each has advantages and disadvantages.

Key Point

Given the ongoing uncertainty around the most appropriate variables, techniques and functional forms of benchmarking models, decisions on model specification generally rely on data availability and opinions on the veracity of results from the various models available, rather than any definitive evidence of the suitability of the model itself.

Key Point

Australia is a particularly challenging jurisdiction for electricity network benchmarking. The small number of networks and large variation in operating conditions and responsibilities of DNSPs requires either more introduction of international data or a more detailed understanding of the cost impacts of environmental variables. Unfortunately meeting one of these requirements often precludes the other.

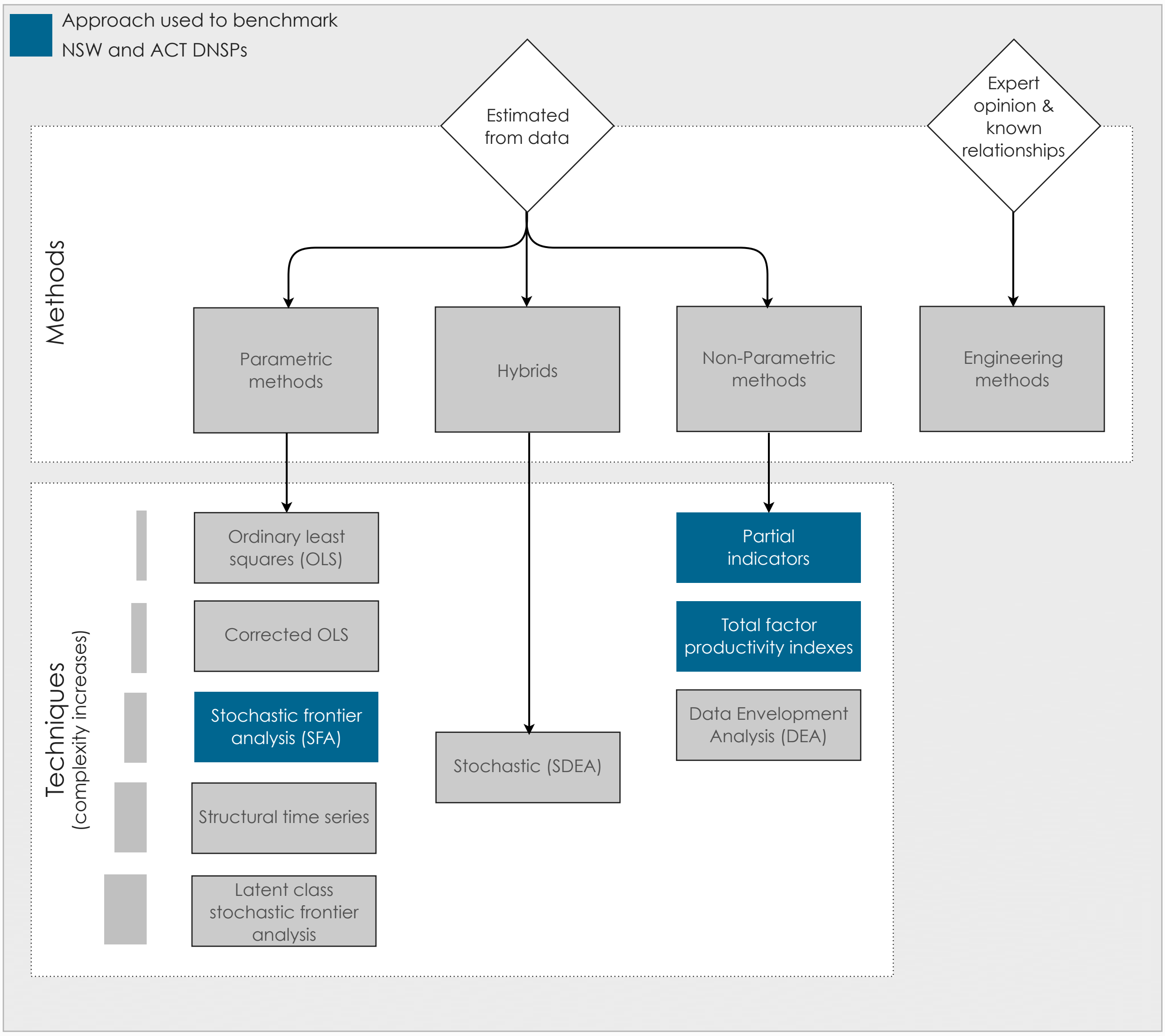
2.1 Benchmarking has many forms

There are many types of benchmarking applied to electricity network cost and performance comparisons and many attributes of each type. For the purpose of this paper, we categorise the attributes of benchmarking in terms of:

1. Benchmarking methods and techniques;
2. Functional specifications; and
3. Model specifications.

Each of these are discussed further in the following sections. The hierarchy of benchmarking methods and techniques is also shown below.

Exhibit 2.1. Benchmarking methods and techniques. *Benchmarking methods are categorised mainly as parametric or non-parametric. These are then broken down further into techniques which describe the basis of combining variables and the reference point of measurement.*



2.1.1 Benchmarking methods and techniques - parametric

Parametric methods rely on assumptions about the distribution of the underlying population to provide estimates of the parameters of a cost function (i.e using mean values and standard distributions to make estimates regarding the underlying population). Common benchmarking techniques that are incorporated within the parametric method include the following.

Ordinary Least Squares (OLS) and Corrected Ordinary Least Squares (COLS)

These techniques seek a line of best fit between the observed data set in order to estimate a relationship between outputs, inputs and environmental variables. Corrected Ordinary Least Squares adjusts this line to form a frontier from which to compare businesses in the sample.

For example, when benchmarking using an industry cost function the line may be shifted down to the firm with the largest negative residual (that is the business that has an actual cost that is the furthest below the predicted cost) and then other firms can be compared to this line. A common criticism of OLS and COLS is that it assumes that the whole distance between the industry line of best fit and a business's actual costs is inefficiency when it could actually result from model misspecification, omitted variables or data errors. The UK regulator, OFGEM, uses this approach to benchmark UK networks. However instead of moving the frontier line to the business that is furthest from the cost line, they use the upper quartile of businesses that appear efficient using the modelling technique.

Stochastic Frontier Analysis (SFA)

Stochastic Frontier Analysis (SFA) is similar to OLS in that it attempts to model the relationship between costs, outputs, inputs and environmental variables. The key difference between SFA and OLS is that it estimates DNSP efficiency by separating the error term (which is assumed to be inefficiency in COLS) into an inefficiency component and a random error component. Whilst SFA is favoured to OLS and COLS because of this adjustment to the error term, it requires a much larger data set and requires additional assumptions about the distribution of the inefficiency term. Regulators in Sweden, Germany and Finland have used this approach in conjunction with other approaches such as DEA. These jurisdictions have many more networks than Australia and benchmark total expenditure.

Latent Class Stochastic Frontier Analysis (SFA)

Latent class SFA builds on the SFA technique but assumes that firms within an industry are likely to operate using different technologies (different parameter estimates) and should therefore be split into different groups for the purposes of benchmarking. These groups can then be used to benchmark DNSPs against other DNSPs within their group (an example in the Australian context might be groupings based on rural and urban businesses). Latent class SFA has the advantage that it allows for greater heterogeneity between businesses and therefore is likely to provide more accurate estimates of efficiency between businesses. A disadvantage of this approach is that it requires a large amount of data to both identify the different groups accurately and perform SFA on these different groups. To date, latent class SFA modelling of electricity distribution networks has been explored through academic channels more than regulatory; however the utility of this extension of SFA models has been found to be high[2].

Structural Time Series

Structural time series models are regression models that incorporate greater flexibility in terms of modelling cost or production functions as they allow for parameter estimates that can vary over time. That is, the regression model parameters are functions of time rather than constants.

2.1.2 Benchmarking methods and techniques - non-parametric

Non-parametric methods do not require the same assumptions regarding the underlying distribution of the population and therefore require smaller amounts of data for use in benchmarking analysis. Common non-parametric techniques are discussed below.

Partial Indicators

Partial indicators are the ratio of a single input and output, for example opex/customer. Whilst this technique is relatively simple in comparison to other benchmarking techniques it does not account for the different factors beyond the control of businesses that influence the ratios.

Total Factor Productivity (TFP)

Total factor productivity incorporates multiple outputs and inputs by using different weights derived from revenue and cost shares to aggregate them into a single output and input index. Total factor productivity is generally preferred to partial indicators because it is

---

2 See, for example: Agrell, Farsi, Fillippini & Koller, "Unobserved heterogeneous effects in the cost efficiency analysis of electricity distribution systems", January 2013 and Llorca, Orea & Pollitt, "Using the latent class approach to cluster firms in benchmarking: An application to the US electricity transmission industry", March 2014

able to include more outputs and inputs through which to benchmark businesses, a criticism is that it is unable to account for environmental differences that can influence the productivity results. TFP has been utilised in electricity network regulation in New Zealand, Canada and the United States.

Data Envelopment Analysis (DEA)

Data Envelopment Analysis uses linear programming to compare businesses with others in its industry. Essentially DEA takes components of multiple businesses in the industry to build a hypothetical firm that produces the same amount of outputs with less inputs (or more outputs for the same amount of inputs). Businesses are then compared to this hypothetical firm and efficiency estimates obtained. Similarly to partial indicators and TFP, DEA uses the combination of outputs and inputs to determine the relative efficiency between businesses. This means that if there are some businesses that have environmental factors that affect their ability to convert inputs into outputs they are likely to appear inefficient compared to businesses operating in more favourable environments. OFGEM has utlised DEA in previous regulatory frameworks, however it is often only used in support of other techniques.

2.1.3 Benchmarking methods and techniques - engineering methods

Engineering methods rely upon the expertise and knowledge of electricity network maintenance and operations to construct a benchmark cost associated with the physical and environmental attributes of a particular network. There are no specific techniques associated with this method, but partial indicators are often used as the unit cost of certain activities that are then extrapolated to the scale and rate considered appropriate to the network characteristics.

All of the methods and techniques described in the preceding sections attempt to identify the relationship between cost and outputs produced - this they have in common. The primary difference between them is the manner in which actual costs are compared to the efficient cost determined by the model. The differences between the techniques in terms of the way in which efficiency is measured is illustrated in exhibit 2.2 on the following page. As shown, the choice of technique dictates the distance from the reference point of efficient expenditure. This highlights the issue of applying these techniques deterministically, as the choice of technique has influence on the assumption of the magnitude of inefficiency.

2.1.4 Functional specifications

When we refer to functional specifications in this document we refer to the functional form of the production function used in econometric models. Econometric modelling is a parametric approach used to estimate the relationship between inputs and outputs. In the context of modelling an opex cost function, modelling will attempt to identify the relationship between operating costs and different cost drivers. A benefit of using econometric modelling to estimate an industry cost function is that it can produce statistical results that can be used to infer which variables have a significant effect on operating expenditure and how well the proposed model explains variation in DNSP expenditure. When modelling operating costs, the relationship between costs and cost drivers needs to be assumed. The two most common functional forms are the Cobb-Douglas cost function and the Translog cost function which are displayed below.

Cobb-Douglas

$$\ln Opex = b_0 + b_1 \ln Y_1 + b_2 \ln Y_2 + b_3 \ln X_k + b_4 T + b_5 \ln Z_1$$

In this example of a Cobb-Douglas model, there are two outputs (Y1 and Y2), one measure of capital (Xk), a time variable (T) and an environmental variable (Z1).

Translog

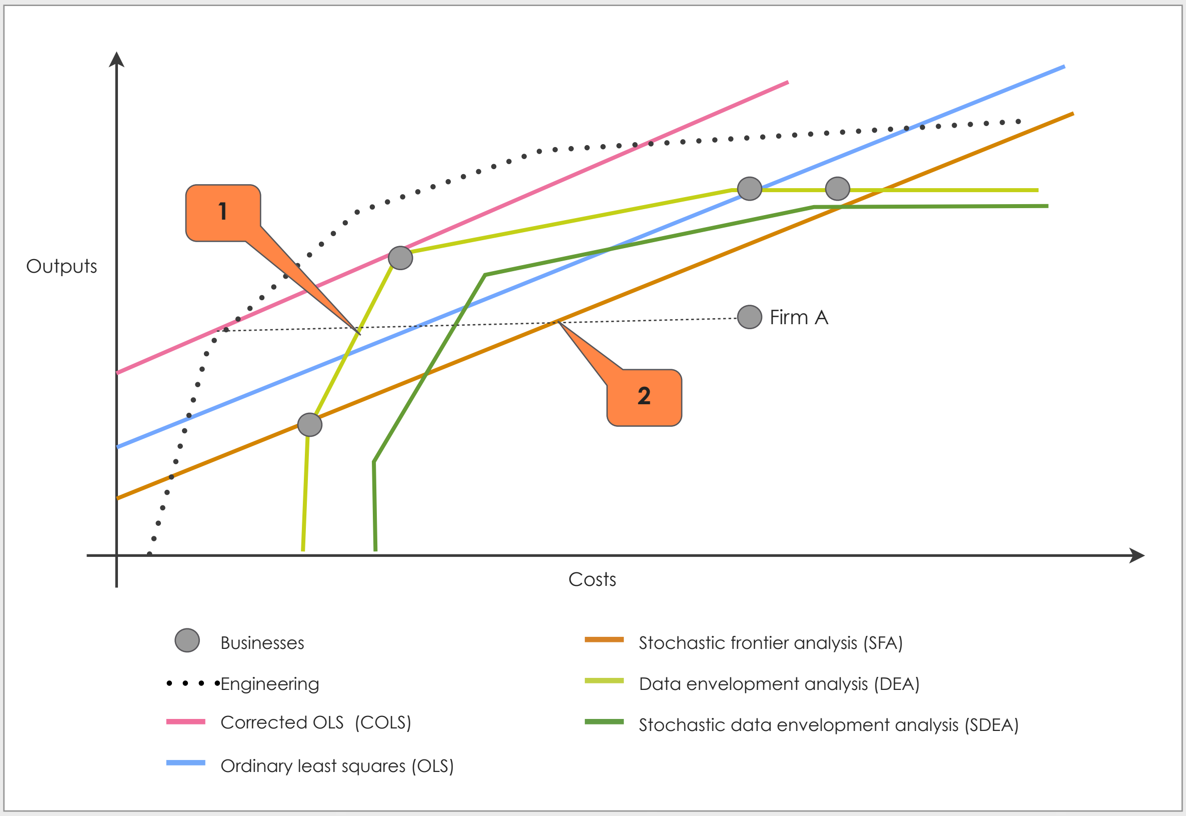
$$\ln C = b_0 + b_1 \ln Y + b_2 \ln X + 0.5 b_3 \ln Y \ln Y + 0.5 b_4 \ln X \ln X + b_5 \ln X \ln Y + b_6 T$$

The Translog model builds upon the Cobb-Douglas model by introducing interaction terms into the model - in this case lnXlnY is included as well as squared terms for outputs (Y) and an inputs (X). The Translog functional form is often preferred to the Cobb-Douglas because it imposes less restrictions on the production and substitution elasticities (i.e one does not have to assume constant returns to scale as imposed with a Cobb-Douglas functional form). One difficulty with using the Translog functional form, particularly within the context of electricity distribution is that electricity distribution tends to have highly correlated variables that exhibit little intra-group variation. This is known as multicollinearity, and it can result in unstable estimates that can change significantly given minor changes in the model specification or underlying data.

2.1.5 Model specifications

When we refer to model specifications in this document we are simply referring to the variables chosen as inputs and outputs. There are vast amounts of literature dedicated to the subject of input and output specifications for electricity networks. This literature not only portrays the difficulty analysts face when attempting to conceptualise what an electricity network actually produces, it also demonstrates the lack of consensus regarding the classification of individual attributes as either an input or an output. For example, many models assume line length as an input and many consider it an output. The ongoing academic conversation around this topic illustrates the challenge in applying techniques suited to production scenarios where the outputs can be both defined and measured (bank transactions, airline passenger miles, products from a factory, patients treated, etc) to the electricity distribution scenario where products delivered are not so easily counted, let alone identified.

Exhibit 2.2: Does the choice of technique matter?



The common characteristic of the approaches listed in the graph above is that they all attempt to estimate the costs a business should incur given the outputs they produce and the environments they operate within. If these methods all gave similar results, then the choice of benchmarking model would be irrelevant; however the graph above highlights the different outcomes that can occur using different techniques, even though they may rely on the same or a similar dataset.

Using Firm A, efficiency scores are obtained by measuring the distance between its actual cost and estimated cost. For example, if we were benchmarking using DEA, Firm A would be compared to point 1, whereas using SFA, Firm A would be compared to point 2. Because Firm A is further away from the DEA frontier it will appear more inefficient using this approach. This graph, whilst only an example, illustrates why different models are likely to provide different estimates of efficiency. As pointed out by Growitsch* there is not one preferred technique to benchmarking electricity distributors:

"Although, from a technical point of view, distribution networks can be regarded as relatively simple activities, there is no consensus in the academic literature or among the regulatory practitioners as how to model this activity"

With this in mind, the approach that appears to be most common in regulatory jurisdictions around the world is to use a combination of results from different benchmarking techniques to arrive at relative levels of efficiency between businesses.

*Growitsch, Jamasb, Wetzel, "Efficiency Effects of Quality of Service and Environmental Factors: Experience from Norwegian Electricity Distribution" 2010

In its Expenditure Forecast Assessment Guideline of November 2013, the AER stated that it may employ:

1. Multilateral Total Factor Productivity (MTFP) to test overall efficiency;
2. Data Envelopment Analysis (DEA) as a cross check of the MTFP results; and
3. Econometric modelling to predict efficient levels of operating expenditure.

We note that in the draft decision for NSW and the ACT, the AER has relied upon MTFP and econometric modelling (the SFA model), but not DEA. It has also used a modified form of the MTFP model to calculate an alternative opex benchmark - using opex as the only input. This is a partial factor productivity, MPFP, rather than total factor model, and it isolates the productivity score to an opex productivity score.

Each of the methods and techniques have benefits and limitations. A comparison of the advantages and disadvantages of the models that the AER has relied upon (Multilateral Total Factor Productivity and SFA) to predict efficient opex and those of the referenced but unused technique (DEA) is shown below.

Exhibit 2.3: Advantages and disadvantages. *Employing multiple techniques can sometimes overcome the individual disadvantages of an individual technique, however data limitations often impede the ability to employ more than one.*

	Total Factor Productivity	Econometric modelling (SFA)	Data Envelopment Analysis
Advantages	An industry cost function does not need to be assumed DNSPs are directly compared to others within the industry and not a regression line (econometric modelling) or a hypothetical frontier business (DEA) The amount of data required is less exhaustive than for other benchmarking techniques MTFP benchmarking is transparent and easy to replicate	Estimates the relationship between different cost drivers and operational expenditure Produces statistical results that can be used to infer which variables have a significant effect on DNSP expenditure and how well the proposed model explains variations in DNSP expenditure Accounts for statistical noise by separating the error term into an inefficiency and a random error component	Weights do not need to be arbitrarily assigned to inputs and outputs Assumptions do not need to be made about the relationship between inputs and outputs of a business The amount of data required is less exhaustive than for other benchmarking techniques such as SFA
Disadvantages	MTFP does not take into account environmental variables, making it difficult to distinguish between inefficiency and the result of different operating environments MTFP does not take into account economies of scale, making it difficult to distinguish between inefficiency and the result of scale differences MTFP scores can change significantly depending on the choice of inputs and outputs MTFP does not produce any statistical results which makes it difficult to determine if the results are valid	Requires more data than DEA and MTFP In the presence of multicollinearity coefficients can be unstable A relationship between inputs and operational expenditure needs to be assumed With a wide range of functional forms and input variables to choose from there may be a number of different models that are statistically valid but produce different estimates	DEA is sensitive to outliers The lack of statistical results means it is difficult to say which variables should be included or omitted DEA results can change significantly depending on which inputs and outputs are being used Businesses will appear more efficient as variables are added

2.2 There is no consensus on the “right” methods and models

Just as there is no consensus on the appropriate inputs and outputs of the model specification, there is no consensus on the most appropriate technique to use for benchmarking DSNPs. Each technique will provide different answers, and often selection of combinations of method, technique and model specification is driven by the available data and other constraints. Whilst some level of statistical testing is available to test the mathematical veracity of some of the techniques, the fitness for purpose of a particular technique and model specification combination often comes down to opinion and judgement of the practitioner. For this reason, regulators and academics often conclude that multiple techniques should be employed. Further, engineering judgement of the veracity of the models and results should always be considered.

Throughout this report we will demonstrate that limited consideration of technique and model specification has been undertaken by the AER and its consultants, which has occurred either through:

1. Rejection of models that produced results considered biased by the AER’s consultants based on scale or location (where other regulators employ multiple models to balance out the bias of an individual model) - see section 4.1.2; or
2. The lack of data available (due to the introduction of a limited set of international data to satisfy the requirements of SFA modelling) for appropriate sensitivity testing.

The significance of such a limited approach is high, particularly if there is reliance upon the results deterministically. The Productivity Commission also recognised the ongoing failure to reach consensus on a single model or approach, further highlighting the need for a balanced view.

> **“** ***The literature on benchmarking is confused. There are:***
>
> - ***multiple methods for benchmarking, with little consensus about which is best; and***
> - ***divergent views about the appropriate inputs and outputs of electricity network businesses.***
>
> \- Productivity Commission, p147 Vol 1 “Electricity Network Regulatory Frameworks”, 26 June 2013

2.2.1 Comparison of Ontario benchmarking results highlights the dangers in relying upon efficiency scores deterministically

The AER has relied upon the analysis of their consultants, Economic Insights, in determining the efficiency of the NSW and ACT businesses in the draft decision. Economic Insights have used a Cobb-Douglas form of Stochastic Frontier Analysis as their preferred technique for measuring efficiency of the businesses. As discussed in this chapter, SFA requires a large dataset and Economic Insights has therefore expanded the dataset to include data from New Zealand networks and Ontario, Canada networks. Issues related to the introduction of data from other jurisdictions (particularly where the imperative is to facilitate a specific model type, rather than any intrinsic value it adds) are discussed later in this report. However to highlight the subjectivity and volatility inherent in the development of efficiency scores, we compared the rankings of the Ontario networks from the Economic Insights model and data[3] with the benchmarking results relied upon by the Ontario Energy Board (OEB). Exhibit 2.4 shows the change in rankings between the OEB’s own benchmarking results and those resultant from Economic Insights’ model of efficiency.

The consultants, models and assumptions utilised by OEB and relied upon by the AER differ, however the dataset is common to both. The OEB analysis is also based on a total expenditure model, whereas the Economic Insights efficiency analysis is based on opex only. We would not expect the two studies to match, however it is reasonsable to expect similar outcomes if benchmarking models are to be used for the purpose of generating a substitute expenditure forecast, with the critical consequences associated with such use. The significant variation in the efficiency rankings of the networks between the two studies highlights the strong dependence of the results on the choices made by the analyst. Given that there is no consensus on the most appropriate model specifications and techniques, this variation also highlights the dangers of placing undue reliance on any one particular set of results when making inferences of efficiency or worst, using the results to determine appropriate adjustments to forecast expenditure. We also note that in the Economic Insights results, the highest and lowest ranked networks have a common, government owner.

---

[3] Huegin used the Economic Insights SFA CD model and data supplied by the AER to reproduce the results.

Exhibit 2.4: Ontario network efficiency variation. *The benchmarking results relied upon by the AER reflect a significant contradiction in efficiency rankings to the views of the Ontario Energy Board.*

Model results: Economic Insights model rankings are sourced from the information released by the AER with the NSW and ACT draft decision. Ontario Energy Board rankings are sourced from the Pacific Economics Group (PEG) analysis at "Empirical Research in Support of Incentive Rate-Setting: 2013 Benchmarking Update - Report to the Ontario Energy Board", July 2014

2.3 Small samples and heterogeneity frustrate benchmarking efforts for Australian networks

Finally, it is worth noting the compounding effect of small samples and heterogeneity on the inherent confusion and debate around model specification and technique. Small samples restrict statistical strength of the analysis and, where small samples are also characterised by a diverse range of operating conditions, any benchmarking technique will be limited in terms of the meaningful information that it can provide regarding the relative efficiencies of the networks in the sample. Large datasets facilitate more robust statistical models and also offer the opportunity to separate clusters of businesses into classes or groups of similar networks. Australia has neither a large dataset nor homogenous conditions.

2.3.1 Augmenting the sample with international data brings its own problems

In the draft decision for the NSW and ACT networks, the AER has relied upon an expanded dataset through the inclusion of New Zealand and Ontario, Canada network data. The Economic Insights model incorporates data from 18 New Zealand and 37 Ontario networks, along with the 13 Australian networks in the NEM. The inclusion of data from New Zealand and Ontario in this regulatory cycle is a decision based on expanding the available dataset to facilitate the introduction of Stochastic Frontier Analysis (SFA) as the econometric modelling technique, rather than precedent used in more mature jurisdictions. International data has been used in academic benchmarking efforts, but it is uncommon that a regulator would use international data in analysis used to set opex[4]. The AER have previously acknowledged the challenges of inclusion of international data.

> “
>
> ***We consider international collaboration of economic benchmarking to be an appropriate goal in the long term and our economic benchmarking should not be limited to a comparison of Australian NSPs. In our view, potential problems with availability of consistent and reliable international data and other analytical issues, may make implementation of an international benchmarking exercise difficult in the short term.***
>
> - AER, Better Regulation, Explanatory Statement, p140 “Expenditure Forecast Assessment Guideline for Electricity Distribution”, November 2014

Extreme caution must be taken when assuming efficacy of the data set and associated modelling results where international data is used. And whilst the data validation process in the AER’s Better Regulation Guideline (see below) may have been intended for the Australian RIN data, the absence of similar prudence for the international data is material given the weight it has (55 of the 68 businesses are international) in determining the SFA model specification and coefficients. We do not consider that the international data meets the Expenditure Forecast Expenditure Guideline’s data validation process.

> “
>
> ***We will commence our data validation process once we have received completed back cast data templates. This process will involve three phases:***
>
> 1. ***We will conduct a preliminary check of data to identify anomalies and correct errors, and a confidentiality check to prepare the data for public release. This will involve bilateral consultation with the relevant NSPs if any issues arise. This is likely to be iterative.***
> 2. ***We will publish refined data to allow interested parties to cross check data and conduct their own analysis.***
> 3. ***Interested parties may provide feedback on overall data quality and any specific data issues that arise.***
>
> - AER, Better Regulation, Explanatory Statement, p162 “Expenditure Forecast Assessment Guideline for Electricity Distribution”, November 2014

[4] The Commission for Energy Regulation in Ireland, with only one distributor has utilised international data for regulation. The Commerce Commission in New Zealand has previously included Australian data for productivity analysis, but not to set opex allowances.

2.3.2 Heterogeneity issues remain

OFGEM is often cited as a source of mature regulatory practice, particularly due to its relatively long history of economic benchmarking for regulatory purposes. OFGEM faces a similar issue to the AER in the small sample size it regulates, however it at least has a more homogenous data set to work with. The spread of network customer densities that OFGEM regulates is shown on the lower left of the graphic at exhibit 2.6 on the next page. The lower right shows, by contrast, the 68 networks (all 13 NEM networks, 37 Ontario networks and 18 New Zealand networks) selected by Economic Insights to include in its SFA model. Great Britain has a similarly small sample size to Australia, however the homogeneity of its networks is much greater (illustrated through the simple measure of customer density). It has many networks of similar size, density and operating conditions, with the exception of the outliers - the London city network and the rural Scottish network. Often, OFGEM will exclude the London and Scottish networks from certain analyses to reduce the influence of outliers (see exhibit 4.2). New Zealand shares similar traits of homogeneity (with density, if not scale) and has one outlier at a higher customer density. The larger land masses of Ontario and Australia produce much greater spreads of network conditions. As shown these businesses span the entire density spectrum. As the Australian data shows, there is an uneven spread of a small sample size over a large range of customer densities.

The structures of the industry in Ontario and New Zealand are of particular concern. These jurisdictions are characterised by one or two large networks and dozens of very small networks. Excluding Hydro One (1.2 million customers, 121,000 km of network and also the worst performing network in the SFA model relied upon by the AER), the remaining 36 Ontario networks included in the model data set have an average network of less than 80,000 customers and around 1,700 km. The Australian data generally has a higher median and covers a higher spread for most common network attributes than the New Zealand and Ontario businesses (see exhibit 2.5 below).

Exhibit 2.5: Sample data distribution. *Australia has a higher median and a greater spread of data for most network measures, other than the percentage of the network underground.*

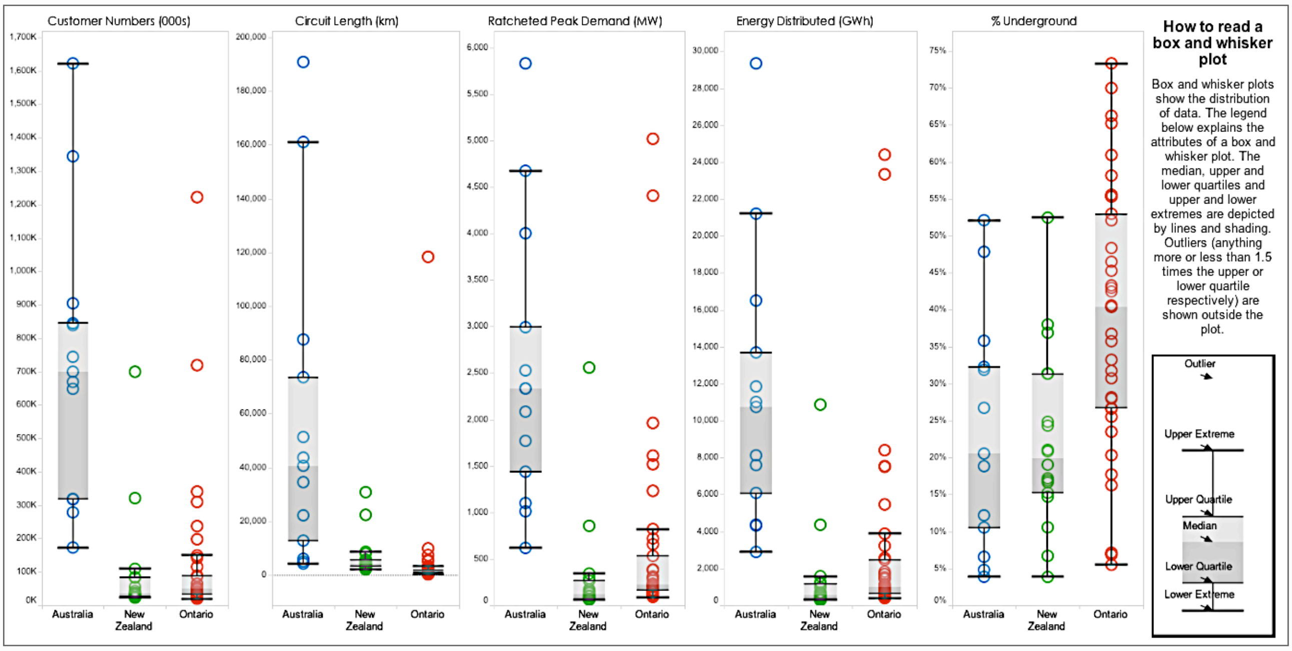


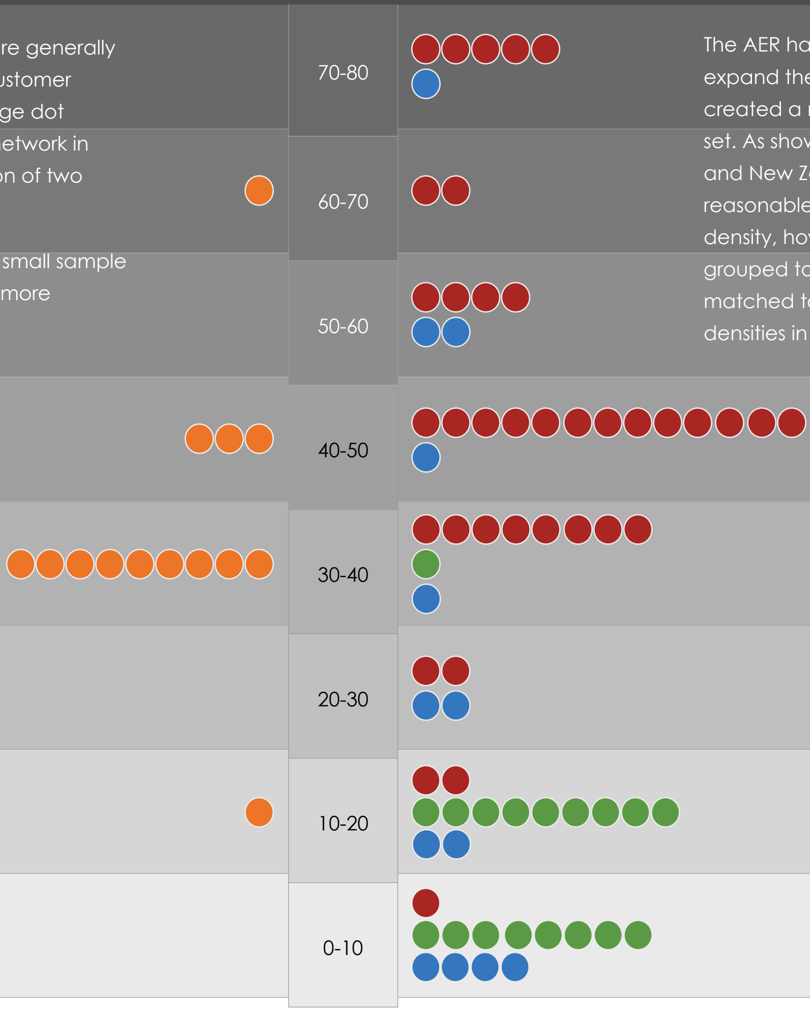
Exhibit 2.6: Compared on Customer Density

Great Britain, Australia (NEM), New Zealand and Ontario

Customer density is number of customers per kilometre of network length

The businesses in Great Britain are generally clustered around a common customer density band (there is one orange dot representing each distribution network in Great Britain), with the exception of two outliers.

Whilst OFGEM are faced with a small sample size, they have the benefit of a more homogenous data set.



The AER have utilised international data to expand the data set, but in doing so have created a much more heterogeneous data set. As shown to the left, the Ontario (red) and New Zealand (green) businesses exhibit reasonable homogeneity of customer density, however they are not similarly grouped to each other and neither is matched to the wide spectrum of customer densities in Australia (blue).

Source: Huegin analysis, public data and Economic Insights econometric data.

Our observations of the AER's benchmarking approach

Our view on the reliability of the benchmarking analysis presented in the AER's draft decision for NSW and ACT is shaped by our opinions on the approach. As such, this section outlines our understanding of the AER's approach to regulatory benchmarking and our associated observations with reference to published best practice and the experiences of other jurisdictions.

3

In this chapter

The following key points are relevant to this chapter on the AER approach

Huegin was asked to review the benchmarking approach for the NSW and ACT draft decision released by the Australian Energy Regulator (AER) in late November 2014. In particular we were asked to address whether the approach leads to results that can be relied upon in setting regulatory allowance for operating expenditure. We have employed an assessment framework outlined in the Productivity Commission's Inquiry Report[5] relating to the regulation of electricity networks to assist with our review. This chapter focuses on our observations of the benchmarking measures used and the agency and statistical practices applied during the process. The next chapter presents our views on the efficacy of the measures and practices.

Key Point

We find the validity and robustness of the benchmark measures used by the AER are limited in their capacity to inform relative efficiency due to:

- the limited explanatory power of the variables chosen; and
- the capacity for the measures to provide signals of efficiency that are more likely statistical noise and error.

Key Point

Robust statistical and agency practices can compensate for limitations of the chosen benchmarking measures to explain efficiency differences. However, we find that aspects of these practices have exacerbated the issues rather than mitigated them - for example, the introduction of international data that facilitates SFA also inhibits robust consideration of environmental variables.

Key Point

Reliance on the benchmarking analysis for regulatory determinations is premature due to the variability of results possible, the lack of consideration of environmental variables, the uncertainty in the data and the immaturity of the approach.

5 Productivity Commission 2013, Electricity Network Regulatory Frameworks, Report No. 62, Canberra.

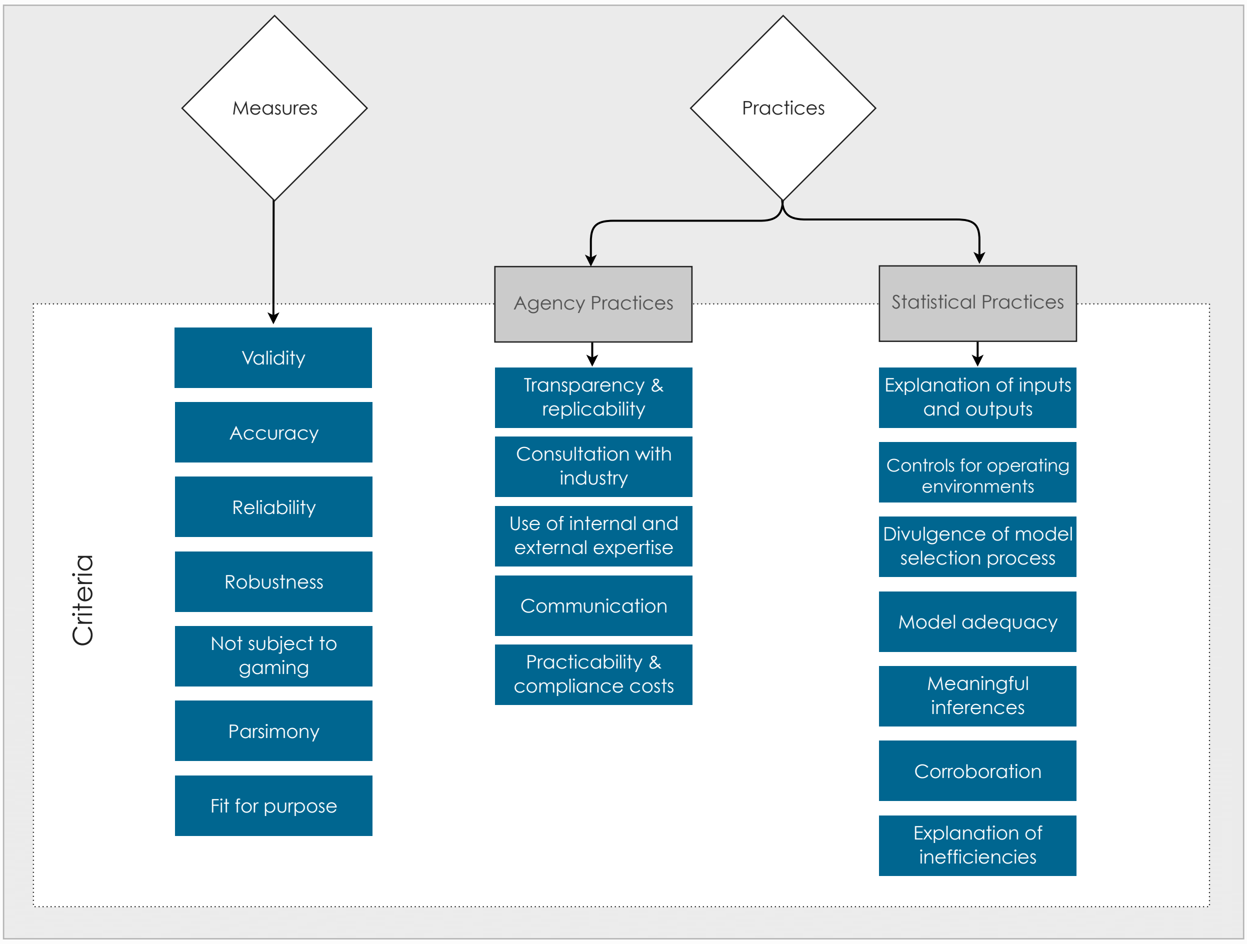
3.1 Defining best practice in regulatory benchmarking

It is useful to have a frame of reference against which to review the benchmarking framework and approach of the AER. We consider best practice can be determined through analysis of the experiences of other jurisdictions and literature on the subject of benchmarking framework implementations. There are several key documents relevant to the AER's benchmarking framework and approach, including:

1. The Productivity Commission Inquiry Report on Electricity Network Regulatory Frameworks[6];
2. The observations from electricity network regulators with more experience in the applied techniques than is currently held by the AER; and
3. The AER's own Expenditure Forecast Assessment Guideline.

These sources are somewhat complementary. The Productivity Commission report and the AER's Guideline and the associated documents that fed into both rely heavily on the experiences of regulators such as OFGEM, the implementation of regulatory benchmarking frameworks around the world and the lessons learnt. For this section we recall the framework within the Productivity Commission report for evaluating the application of network benchmarking, shown below at exhibit 3.1.

Exhibit 3.1: Benchmarking Assessment Framework. *Benchmarking measures and practices can be evaluated against criteria designed to inform the explanatory power of the measure(s), efficacy of the models used to generate them and the appropriateness of the process to deploy and utilise them.*



Adapted from: Figure 4.7 of Productivity Commission 2013, Electricity Network Regulatory Frameworks, Report No. 62, Canberra.

6 Productivity Commission 2013, Electricity Network Regulatory Frameworks, Report No. 62, Canberra.

3.2 Benchmarking measures

Benchmarking measures refer to the quantitative measurement factors used by the AER to infer efficiency differences between the networks. For the purpose of this critique, we consider the relevant benchmarking measures to be:

1. The MTFP efficiency scores (industry, state and individual DNSP) that the AER has relied upon for an indication of relative efficiency;

2. The SFA model efficiency scores that the AER has relied upon to determine the efficient level of base year operating expenditure; and

3. The opex MPFP efficiency scores that the AER has relied upon to validate its other analysis.

The focus of this section is the latter two measures, which are described in exhibit 3.2 below.

Exhibit 3.2: The AER's opex benchmarking techniques. *The SFA and Opex PFP benchmarking measures relied upon by the AER are explained through the model specifications outlined below.*

Stochastic Frontier Analysis

Stochastic Frontier Analysis uses maximum likelihood estimation to model the relationship between a dependent variable (opex), a number of different explanatory variables (circuit length etc.) and the cost efficiency of each unit (DNSP). An efficiency score is obtained by making assumptions regarding the distribution of the error term ($v_i + u_i$) to separate DNSP inefficiency (u_i) and random noise (v_i). Cost efficiency is then calculated as exp(u_i).

$$\ln Opex = \alpha + BX + v_i + u_i$$

The AER's model

The model used by the AER uses customer numbers, circuit length, ratcheted maximum demand, share of underground, year and a dummy variable for Ontario and New Zealand to explain DNSP opex.

$$\ln Opex = \alpha + \ln(CustNum) + \ln(CircLen) + \ln(RMDemand) + \ln(ShareUG) + Year + NZ + Ontario + v_i + u_i$$

v_i has been assumed to be normally distributed
u_i has been assumed to have a truncated normal distribution
ln indicates the variable enters as a natural log

Opex Partial Factor Productivity

Opex PFP uses an index of outputs relative to opex to produce an Opex PFP score that can be compared across DNSPs. Weights are used to from an aggregate output index from a number of different outputs. The technique compares output levels from a reference firm to ensure output indexes can be compared across DNSPs and time.

$$\ln(Y_{st}^{T*}) = \begin{bmatrix} \sum_{i=1}^{N} \left(\frac{\omega_{it} + \overline{\omega}_i}{2} \right) (\ln y_{it} - \overline{\ln y_i}) - \\ \sum_{i=1}^{N} \left(\frac{\omega_{is} + \overline{\omega}_i}{2} \right) (\ln y_{is} - \overline{\ln y_i}) \end{bmatrix}$$

Where;

ω_i is the output weight (a bar above the variable indicates the industry average)

y_{it} is an output

The second row of the formula is the calculation for the output index of the reference firm. Note that an input index is calculated using the same formula but with a single input variable (opex).

The AER's model

The AER has used the following outputs in the construction of an output index, weights are included in brackets.

Energy distributed (0.128), Ratcheted Maximum demand (0.176), Customer numbers (0.458), Circuit length (0.238), Total customer minutes off supply (a negative weighting calculated using AEMO's Value of Customer Reliability)

The benchmarking measures used can be assessed against their ability to adequately explain variations in costs and efficiency between businesses in the benchmarking group. There is no perfect measure of electricity network benchmark cost, but some are better than others. It is also important to note that a less than optimal benchmarking measure may be compensated by the statistical practices applied. For example, a benchmarking measure that is not robust due to sensitivity to changes in operating environments can be made useful through the consideration of environmental variables in the statistical practice of applying the measure.

Our views on the benchmarking measures employed by the AER in both the annual benchmarking report and, more substantially, in the draft decision for NSW and ACT are outlined in the following paragraphs which follow the categories identified in exhibit 3.1.

3.2.1 Validity

Validity is defined as the extent to which the benchmark actually measures the attribute under study - in this case, efficiency. The benchmark measures relied upon by the AER compare businesses on the basis of customer numbers, line length and ratcheted peak demand. Any residual between an individual business and the business(es) that are deemed to have the most efficient relationship between these variables and expenditure is composed of:

1. The influence of variables that are not included in the model yet influence costs;
2. Measurement error; and
3. Efficiency differences.

Without correcting for the first two, the residual is not a valid measure of efficiency. This is not always an issue so long as the measures are robust and the statistical practices sound (and that the results are not used inappropriately).

Our primary concern with the validity of the opex SFA and MPFP efficiency measures is the assumption that operating expenditure can be explained through customer numbers, ratcheted peak demand and line length. In our experience and supported by previous Huegin studies[7] , the incremental opex cost of increasing customers in an electricity network is actually quite low, yet the SFA model has a coefficient[8] of 0.667. The incremental opex costs of increasing ratcheted peak demand is negligible, yet the SFA model gives this variable a coefficient of 0.21. Line length is the only variable that presents some sort of proxy for the asset itself, but even then:

1. It's influence is low in the model, with a coefficient of only 0.10; and
2. The actual line length is only a moderately strong proxy for influence of the asset on opex, as the design, type and location of the assets all drive opex.

Overall, we consider that the validity of the model is poor. Whilst there are coincidental relationships between increases in customers and line length, more important considerations of the asset design and location are not considered.

3.2.2 Accuracy

The accuracy of benchmarking measures relates to the ability to provide an unbiased estimate of efficiency. Because of the one-size-fits all structure of the SFA formula, with its coefficients fixed for all businesses, we do not believe that the measure is an accurate reflection of efficiency. The reasons for this are explored further in chapter 4 of this report.

The use of averages over time also erode the accuracy of the measures as they relate to current efficient expenditure, as businesses that have increased expenditure will reap the benefit of the historical average against businesses that have not. Taking historical averages also assumes that past conditions are representative of current and future conditions, which is unlikely given the amount of change and reform in the industry over the past four years. Section 4.3.1 presents data that demonstrates the dynamic nature of the industry costs over time and the consequence of using historical averages.

Data normalisation is another important consideration for result accuracy that has not been appropriately reflected in the measures. The businesses have significantly different accounting structures and are subject to costs that others are not. OFGEM realise the importance of adjusting expenditure to a comparable basis and make adjustments to the following four categories[9]:

1. Regional labour costs;
2. Company specific factors;

[7] See section 4.1.3 of this report for an overview of cost categories and the associated opex drivers.

[8] A coefficient is the numerical factor that the variable is muliplied by in the algebraic equation of the econometric model. A customer number coefficient of 0.667 for example, means that for every one per cent increase in customer numbers, a 0.667% increase in opex occurs.

[9] OFGEM, "RIIO-ED1: Final determinations for the slow-track electricity distribution companies", November 2014

3. The following exclusions from totex models:
    1. Transmission connection point (TCP) charges
    2. Critical national infrastructure (CNI)
    3. Rising and lateral mains (RLM)
    4. Improved resilience
    5. Quality of service (QoS)
    6. Smart meter roll out (including smart meter call out costs)
    7. New streetwork costs.
4. Other adjustments.

> ***We consider whether DNO submitted data require adjustments prior to carrying out our comparative benchmarking. This is to ensure the comparisons are on a like for like basis. Where we decide adjustments are appropriate, we adjust the DNO submitted costs before our totex and disaggregated assessments. These adjustments fall into four broad categories.***
>
> - OFGEM, p41. "RIIO-ED1: Final determinations for the slow-track electricity distribution companies", November 2014

This is a significant difference in the approach of OFGEM and the AER. The OFGEM approach is also based on many years of regulatory reporting to a consistent format and common reporting timeframes (i.e. the lack of a staggered reporting and/or regulatory determination cycle, as exists in Australia), which are more favourable conditions for data accuracy. Yet OFGEM still recognise the need to normalise the data prior to modelling. Regional and company specific factor adjustments recognise that particular locations and particular networks incur costs beyond the control of the operating business and these costs should not be included in efficiency models. Just as important, the totex model exclusions (point 3 above) exclude costs on the basis that these costs are not adequately explained by the model variables or have changed over time.

> ***These are costs that are inappropriate for comparative benchmarking because they are not adequately explained by cost drivers that are being used in the totex models or because there is a substantial change in the nature of the activity between DPCR5 and RIIO-ED1.***
>
> - OFGEM, p41. "RIIO-ED1: Final determinations for the slow-track electricity distribution companies", November 2014

There are many costs in the operating expenditure of Australian DNSPs that fall into a similar category of inability to be explained by the SFA and MPFP models. Exhibit 4.3 in this report demonstrates the lack of explanatory power of the SFA model variables for many categories of cost. Vegetation management, for example, is a material expenditure category that is neither appropriately described by the model variables nor has it remained a constant activity over time. We note that the AER have made allowances for some of these types of anomalies by adjusting the outputs of the models, however we consider that:

1. It is more appropriate to adjust the input costs, than attempt to correct the output results;
2. There has been insufficient consideration of all cost differences across all networks in the study; and
3. There has been no recognition of cost exclusions based on the failure of the model to describe those costs.

The AER now relies upon Regulatory Information Notices (RIN) that have more commonality than previous reporting templates, however:

1. This is the first year of data reporting, and as such the regime is not mature;
2. The Australian DNSPs still report based on their own accounting systems, and the basis of preparation documents associated with the RINs demonstrate that data consistency and comparability is still an issue; and
3. The AER will always be challenged by data issues due to the lack of common reporting timeframes and the existence of a staggered regulatory cycle.

As an example of the issues related to the third point above, the NSW and ACT draft decision makes some allowance for anomalies in the cost data (OH&S requirements, bushfire obligations, etc), but without adjusting the data for *all* businesses in the sample, the reference point of the econometric models efficiency assessment (the frontier) is unreliable.

3.2.3 Reliability

Reliability is defined as the extent to which it the benchmarking model employed can reliably reproduce a result. The benchmarking measures (SFA and MPFP efficiency scores) employed by the AER's consultants and relied upon by the AER are reproducible with the current model and data provided. However, we question whether the ability to reproduce results will endure over time given the reliance on data from international jurisdictions which may or may not continue to publish the same data in the same format. We are particularly concerned with the reliance on data from Ontario, where restructuring and network mergers are currently under consideration[10].

3.2.4 Robustness

The robustness of the benchmarking measure relates to its ability to produce information about the efficiency of the business regardless of its operating environment. That is, robustness measures the extent to which it is portable across businesses in different environments. In the case of the NSW and ACT draft decision, robustness is particularly important due to the introduction of data from New Zealand and Ontario. The likelihood of these jurisdictions having similar geographic or legislative and policy environments is very low and therefore raises the probability that the efficiency results will be subject to statistical noise and error.

Percentage of the network underground is the only environmental variable considered in the econometric model, because it is one of the few available across all three jurisdictions of the source data. The convenience of data availability does not result in a sound benchmarking model and the absence of more appropriate operating environment data highlights the probability that the model is subject to omitted variable bias and should not be used deterministically.

Furthermore, the model specification itself has evolved from a process of trial and error, with the rejection of alternative model specifications based on observations of the placement of urban and rural or small and large networks in the efficiency rankings.

> “ ***The results obtained using output specification #4 did not appear to favour any particular type of DNSP with both rural and urban, and small and large DNSPs interspersed. Along with its superior in principle characteristics, this lent further support to using output specification #4 as the preferred specification.***
>
> - Economic Insights, p41. “Memorandum on DNSP MTFP Results”, 25 July 2014

Huegin has two concerns with this process:

1. Dozens of input and output variables have been individually supported or considered superior “in principle” throughout the evolving model specification process between the AER's Expenditure Forecast Assessment Guideline and the NSW and ACT draft decision, as they have in the broader international benchmarking environment. However they are quickly abandoned when the models do not behave as expected; and
2. Urbanity and scale are only two of many categorical attributes of electricity networks. In the draft annual benchmarking report (via the supporting memorandum, referenced above), models were discarded if they appeared to favour urban, rural, large or small networks. Apart from the subjectivity inherent in this exercise, such attributes are overly simple and a narrow view of the categorical differences between networks.

We note that the AER also relies upon the observations of relative rankings of urban, rural, large or small networks to determine the robustness of the economic benchmarking models and results.

> “ ***The results of our benchmarking have shown an even spread of results across different types of distributors. The two most efficient businesses on average over the 2006–2013 period are urban and rural networks respectively. The two least efficient businesses on average over the 2006–2013 period are the smallest and largest (in terms of customer numbers).***
>
> - AER, p46, Electricity Distribution Network Service Providers “Annual Benchmarking Report”, November 2014

[10] Ontario Distribution Sector Review Panel, “Renewing Ontario's Electricity Distribution Sector: Putting the Consumer First”, December 2012

Mixtures of businesses at either end of the efficiency rankings based on these two categories (urbanity and scale) is not compelling evidence of the absence of bias, just as coincidence of those network types at either end of the rankings is not sufficient evidence of bias. The issues with such a coarse, limited test of the efficacy of model results include:

1. The classification of a network as either urban or rural is based on its customer density. In Victoria, urban businesses (CitiPower, Jemena, United Energy) operate exclusively within the greater Melbourne metropolitan area and the rural businesses (Powercor and AusNet Services) operate exclusively in the regional and rural areas of Victoria. In South Australia, SA Power Networks operates the entire state network, and because its rural component has a greater line length than its urban component, it is classified by its customer density as rural. Ausgrid and Endeavour Energy, on the other hand, have networks in the Sydney metropolitan area, major regional centres and isolated, rural areas in between. But because of the greater weight of population in Sydney, they are considered urban. Similarly the classification of scale is dependent on the variable used - line length, customer numbers, service area, energy, etc.

2. There are many more determinants of bias in the comparison of electricity network costs that have not been considered. Many of these other genuine differences between networks can be observed at either end of the rankings of the econometric model relied upon by the AER in its NSW and ACT draft decision. Using the same logic as that used by the AER in rejecting and accepting model results, we could hypothesise that the current model presents bias on the basis of network design, population dispersion and climate.

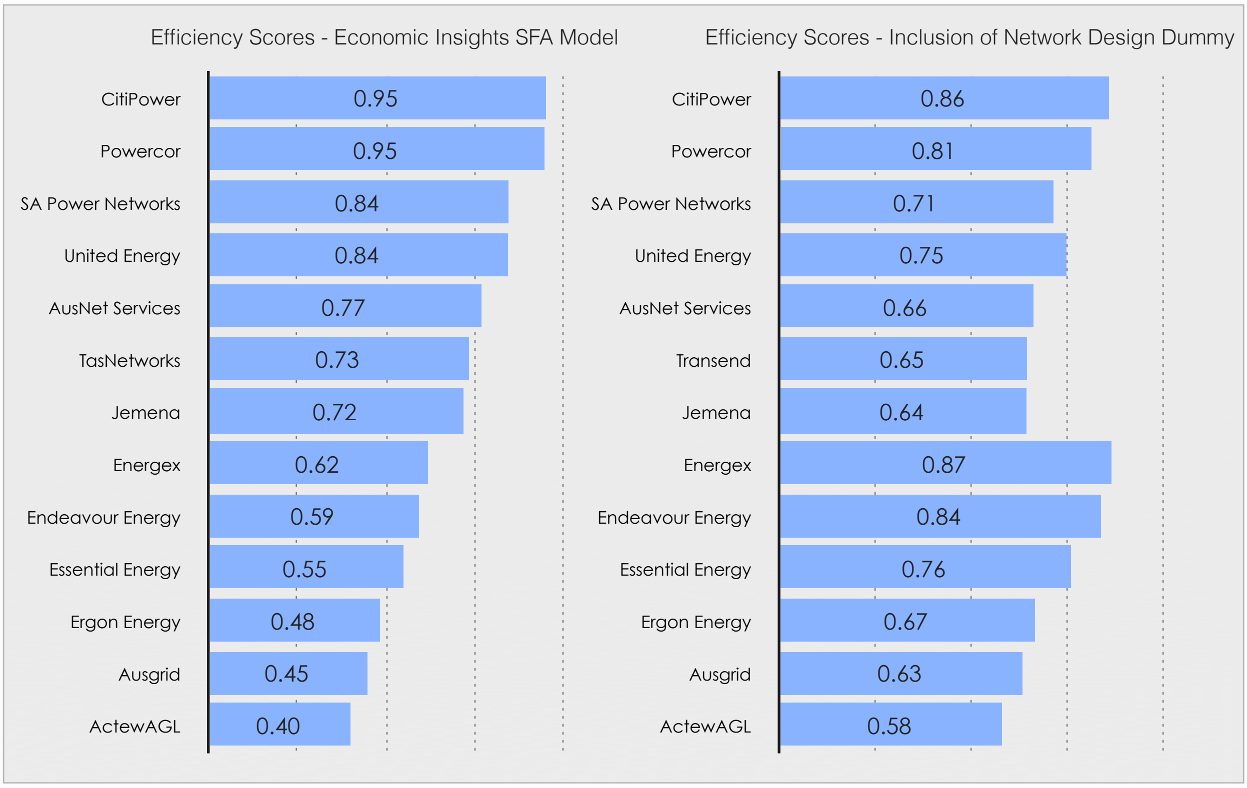
The section on environmental variables in chapter 4 presents analysis on the three factors mentioned above at point 2. However by way of example we considered the network design hypothesis that the current SFA model relied upon by the AER presents bias against networks with high voltage subtransmission assets. We based this hypothesis on the premise that:

1. The bottom six ranked Australian businesses (ActewAGL, Ausgrid, Essential Energy, Ergon Energy, Energex and Endeavour) in the AER's results have assets above 66kV above whilst the top ranked seven have none; and

2. Five of the bottom six ranked businesses in the AER's results have multiple stage transformation within their networks, whereas six of the top seven have solely single stage and the seventh (SA Power Networks) has only a small share (less than 3%) of multiple stage transformation.

If we considered that our hypothesis was valid, we could re-model the SFA econometric model using a dummy variable for the existence of these assets. Exhibit 3.3 shows the original results for the Australian businesses from the Economic Insights SFA model relied upon by the AER and the change in those results if the network design dummy variable is included in the model.

Exhibit 3.3: Model sensitivity to environmental variable selection. *Inclusion of network design factors to account for the existence of assets above 66kV in some networks highlights the dependence of the results on the assumptions made.*



We note that the AER has inferred that the efficiency results are state based[11], but we also observe that our classification of networks through the existence of high voltage assets falls along the same state lines. This highlights the point that casual observations of differences between networks can significantly impact the interpretation of the results. This can potentially lead to erroneous conclusions of efficiency which may in fact be influenced by other variables or network attributes not observed or considered.

3.2.5 Not subject to gaming

It would be difficult to game the model given the variables are generally outside the control of the business to change (customers, line length and peak demand). One issue the AER may have to address in using an econometric cost function to benchmark opex whilst using physical inputs to measure its MTFP is the incentive to shift costs from opex to capex.

There is an inherent benefit in the econometric SFA model to shift operating costs into capitalised costs. Currently, businesses with a high opex to capex ratio (such as ActewAGL) are penalised by the inclusion of opex only in the efficiency estimates of the SFA model. We note that:

1. The best performing network (Hydro One Brampton) amongst the entire sample of the SFA model has one of the lowest opex to capex ratios (27%) in the sample; and
2. The worst performing network (ActewAGL) amongst the entire sample of the SFA model has one of the highest opex to capex ratios (95%) in the sample.

Whilst currently the relationship between levels of capitalisation and efficiency performance under the SFA model relied upon by the AER is circumstantial, over time the incentive to shift costs to capital could be an unintended consequence of the AER's use of the SFA model for opex efficiency evaluation. There are suggestions that this has occurred previously in Ontario due to the focus on opex efficiency and absence of proportionate penalties for reliability under-performance[12 13]. We note that both the Ontario Energy Board and OFGEM have progressed to total expenditure benchmarking.

3.2.6 Parsimony

Parsimony relates to the level to which complexity has been added without adding commensurate value to the analysis, with value measured by increased explanatory power of the model. Econometric analysis can be complex by nature, particularly for practitioners unfamiliar with its use. It could be suggested that, with the heavy weighting on customers as an opex explanatory variable, that similar results could have been obtained using simple regressions of opex per customer ratios.

3.2.7 Fit for purpose

If the SFA and MPFP benchmarking measures were used to explore the reasons for such a large variation of efficiency scores across networks, then they may be considered fit for purpose. However there has been insufficient consideration of other variables and environmental factors acting upon the benchmarking results. Therefore we consider that the SFA and MPFP benchmark measures are not fit for the purpose because they appear to have been used to estimate adjustments to base year operating expenditure rather than to explore reasons for apparent inefficiency.

We also consider that an almost complete focus on historical data, rather than the forecast provided by the businesses, undermines the fitness for purpose of the measures as determinants of efficient levels of expenditure in the forthcoming regulatory period.

---

11 AER CEO, Michelle Groves, Transcript of speech at 2014 Annual Energy Users Association of Australia Conference in Melbourne, "We are finding that the most efficient distributors are those located in South Australia and Victoria. Under most of the performance measures we use, the Victorian and South Australian distributors appear more productive in their use of operating and asset costs. The 'productivity gap' is not even that close.", 13 October 2014

12 Cronin, F.J, Motluk, S., Utilities Policy Journal Vol 19, 2011, "Ten years after restructuring: Degraded distribution reliability and regulatory failure in Ontario", 9 July 2011

13 Power Workers' Union, "Draft Report of the Board on Empirical Research to Support Incentive Rate-setting for Ontario Electricity Distributors", 25 September 2013

3.3 Agency practices

It is not within Huegin's field of expertise, nor is it appropriate, for us to comment on many aspects of the agency (AER) practices in applying the benchmarking analysis through the process of the annual benchmarking report and draft decision for NSW and ACT. We can, however, make observations against many of the criteria as a business engaged by many of the DNSPs to assist them in:

1. Their consultations with the AER on benchmarking issues;
2. Drafting of responses to AER issues papers; and
3. Providing analytical support and advice on benchmarking during the regulatory determination process.

Material observations of the agency practices are included against the criteria discussed individually below.

3.3.1 Consultation with industry

Industry was extensively consulted by the AER during the pre-work to the Expenditure Forecast Assessment Guideline with 18 workshops and numerous opportunities to submit formal responses to issues papers and draft guidelines. When the AER deviated away from the Guideline (model specification and techniques) in the draft annual benchmarking report, industry again was afforded the opportunity to provide feedback.

However we consider that the significant changes that have occurred since the Guideline's release are consequential to the NSW and ACT determination and have not been distributed for consultation. We note that:

1. The MTFP specification has been changed twice - firstly the preferred and alternative specifications from the Expenditure Forecast Assessment Guideline were discarded when the results were rejected by Economic Insights and then the specification was modified between the draft annual benchmarking report and the NSW and ACT draft decision; and
2. The techniques of SFA and OLS were not communicated to the businesses as the preferred techniques until they appeared in the supporting documentation of the NSW and ACT draft decision.

The delay in the final benchmarking report has also provided little opportunity for the NSW and ACT businesses to respond other than in the context of the revised regulatory proposal. We consider that the introduction of SFA and the international data associated with it, combined with the level of reliance and deterministic manner in which it has been used, contradicts the AER's own Guideline:

> “
>
> ***We consider stakeholders should be informed of preliminary economic benchmarking results before they are adopted in our draft and final regulatory determinations.***
>
> - AER, Better Regulation, Explanatory Statement, p162, “Expenditure Forecast Assessment Guideline for Electricity Distribution”, November 2014

The graphic on the following page depicts the asymmetric nature of the consultation on the models and specifications that have been discarded against the multiple evolutions of the model in the short timespan between finalisation of the Guideline and the release of the draft decision.

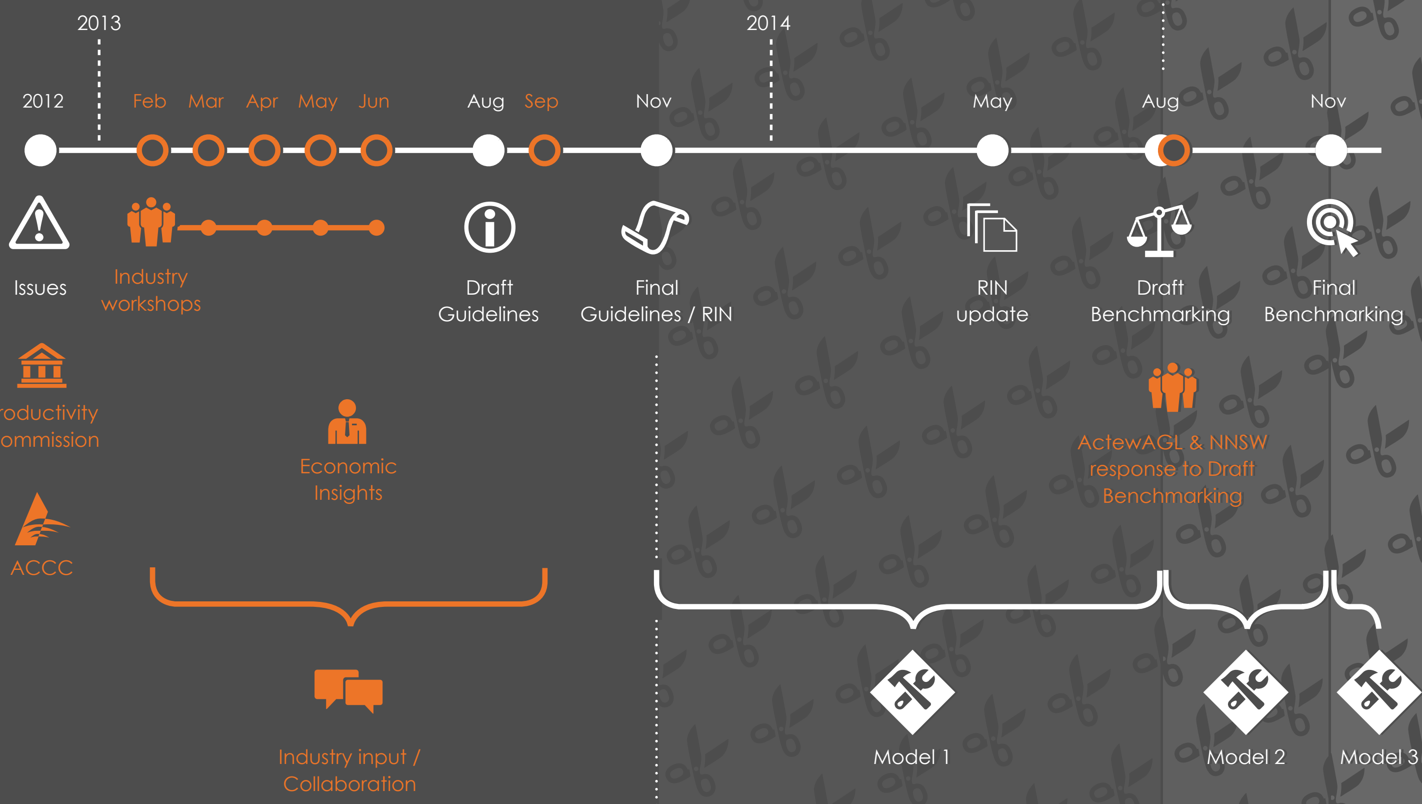
3.3.2 Practicability and compliance costs

We cannot comment on the compliance costs associated with the benchmarking regime or how practical it will be to maintain and manage into the future. We can however offer an observation relevant to the compliance cost of the regime. The large volumes of data collected through the Regulatory Information Notices (RINs), particularly the economic benchmarking RIN has not been used in the benchmarking. This is in part due to the inclusion of New Zealand and Canadian data, for which equivalent data is not available. It is also reflective of the model selection process where experiments with available data and subjective assessments of results have shaped the model specifications to a greater extent than any theoretical merit of the models.

Shifting goalposts

At every stage since the Final Guidelines the AER have changed/altered the benchmarking models based on feedback - how can they be sure they now have the "right" model?

Examination of the results of applying this specification to Australian DNSP panel data revealed that it tended to artificially advantage DNSPs with long line lengths"
Source: Economic Insights



...this shortlist does not exclude other variables from being included in future economic benchmarking analysis as more data becomes available"
Source: AER

3.4 Statistical practices

Statistical practices are of paramount importance in electricity network benchmarking. A simple cost production model will never fully represent the complexities and scope of variables that influence the costs of all businesses in an industry. As such, it is imperative that adjustments for operating environment differences, model specification processes and sensitivity testing is comprehensive and rigorous, otherwise the issues of validity and robustness outlined in section 3.2 will influence the efficiency assumptions.

3.4.1 Explanation of inputs and outputs

The inputs and outputs of the models have been fully explained, quantified and described. Whilst the selection of variables may be debatable, they are clearly defined and explained.

3.4.2 Controls for operating environments

Controlling for operating environments adequately is critical to the success of removing the influence of omitted variables from the residual. This is particularly important given our concerns around the validity and robustness of the MPFP and SFA benchmark measures, We consider that the inability to control for environmental variables is a major shortcoming of the approach. The decision to adopt SFA modelling has necessitated more data, which has been sourced from New Zealand and Ontario.

> ***We thus concluded that to obtain robust and reliable results from an econometric opex cost function analysis we needed to look to add additional cross sectional observations which meant drawing on overseas data, provided largely comparable DNSP data were available.***
>
> - Economic Insights, p29, "Economic Benchmarking Assessment of Operating Expenditure for NSW and ACT Electricity DNSPs", 17 Noveber 2014

This reliance on international data has limited the ability to consider environmental variables to that which is available across all three jurisdictions - share of underground network. Huegin raised the concern of environmental variables and differences in operating conditions in a response to the AER's draft Expenditure Forecast Assessment Guideline. The AER's response was published in the final Guideline Explanatory Statement:

> ***Our broad range of data requirements is designed to allow for rigorous sensitivity analysis in order to test the robustness of our economic benchmarking analysis and to further understand the relationships between inputs, outputs and environmental variables. This will also assist in identifying and correcting for potential shortcomings or econometric issues, such as 'missing-variable bias', in the proposed econometric models.***
>
> - AER, Explanatory Statement, p138, "Better Regulation Explanatory Statement, Expenditure Forecast Assessment Guideline", November 2013

However since the release of the Guideline, the change in modelling technique (Stochastic Frontier Analysis) has constrained such considerations:

> ***With regard to operating environment variables, due to the lack of operating data available for Ontario, we were limited to the inclusion of the share of underground cable length in total line and cable length in this instance.***
>
> - Economic Insights, p32, "Economic Benchmarking of NSW and ACT DNSP Opex", 17 November 2014

3.4.3 Divulgence of model selection process

The process for selecting the model has been divulged by the AER. However it is through that process that issues with the model selection itself have become apparent. The model selection process leading up to the Expenditure Forecast Assessment Guideline was based on literature research and academic arguments. Since the data has been available, the model selection process has relied upon subjective assessments of the veracity of the results. There is a risk that the selected model contains bias of a source not considered. For this reason, most other regulators that rely upon these types of models and approaches include multiple variants and weight them, so as not to rely solely on one particular model specification. The AER has discarded models that do not support its preferred model specification, including an econometric model presented to them by one of their own consultants, Pacific Economics Group, which provides significantly different results to those relied upon by the AER[14]. The lack of consideration of alternate models and weighting of results is of concern.

3.4.4 Model adequacy

Model adequacy extends beyond the statistical significance of the model attributes to a more critical review of the practical usefulness of the model. We cannot find evidence of robust engineering review of the model and its mechanics in the draft decision documentation.

3.4.5 Meaningful inferences

We consider that the models adopted by the AER are capable only of providing inferences of *differences* across the networks, but cannot alone explain efficiency or lack thereof. Having only adopted techniques such as SFA immediately prior to its first use in an electricity distribution regulatory determination, we believe that it is unsafe to rely upon the results of these models to derive adjustments to base year opex.

> “
>
> ***Sweden relied on SFA, together with econometric analysis and DEA, to determine the industry wide productivity change for controllable opex...***
>
> ***Germany and Sweden had over 170 utilities and Sweden also used panel data. Finland had 88 utilities and was not able to include many cost drivers due to this relatively smaller sample size....***
>
> ***It is also important to note that the application of cost benchmarking in each of these three European countries is relatively new. It was introduced in 2008, 2009 and 2012 for Finland, Germany and Sweden respectively. Indeed, it may be premature to draw lessons from these experiences.***
>
> - ACCC/AER, p105, “Working Paper No.6”, May 2012

3.4.6 Corroboration

Corroboration is the extent to which efficiency results (scores, rankings and performance groups) are consistent across alternative specifications and techniques. An often cited technique to test for mutual consistency across model specifications, techniques and functional forms is the application of Bauer consistency criteria. The Bauer consistency criteria were first used to measure consistency of efficiency methods in the regulated financial sector[15]. Since then, they have been cited many times in utility regulation[16]. The first three criteria are focused on measuring the mutual consistency of benchmarking approaches - the level to which different models independently produce reasonably consistent results. The process of testing for mutual consistency involves testing ranges of model specifications and techniques to observe consistency over three levels:

1. Level 1 - consistency of efficiency scores: ideally, different models and techniques will provide consistency of the efficiency scores for businesses. This rarely occurs over a reasonable range of models.

2. Level 2 - consistency of efficiency rankings: if scores are not consistent, but rankings are, analysts can at least make inferences about the relative comparisons of efficient businesses.

---

[14] Pacific Economics Group, “Database for Distribution Network Services in the US and Australia", 21 August 2014

[15] Bauer, P., Berger, A., Gerrier, G. and Humpherey, D., Journal of Economics and Businesses “Consistency conditions for regulatory analysis of financial institutions: a comparison of frontier efficiency methods”, 1998

[16] See for example, Berg, S.V., IWA Publishing, “Water Utility Benchmarking - Measurement, Methodologies and Performance Incentives”, 2010

3. Level 3 - best and worst performers: the final level of consistency test is if the top and bottom groups are consistent then inferences can be made at least about best and worst performers.

Huegin conducted a number of sensitivity tests on the models, techniques and data available and found very little consistency. The results of this analysis are shown in section 4.1.1 of this report.

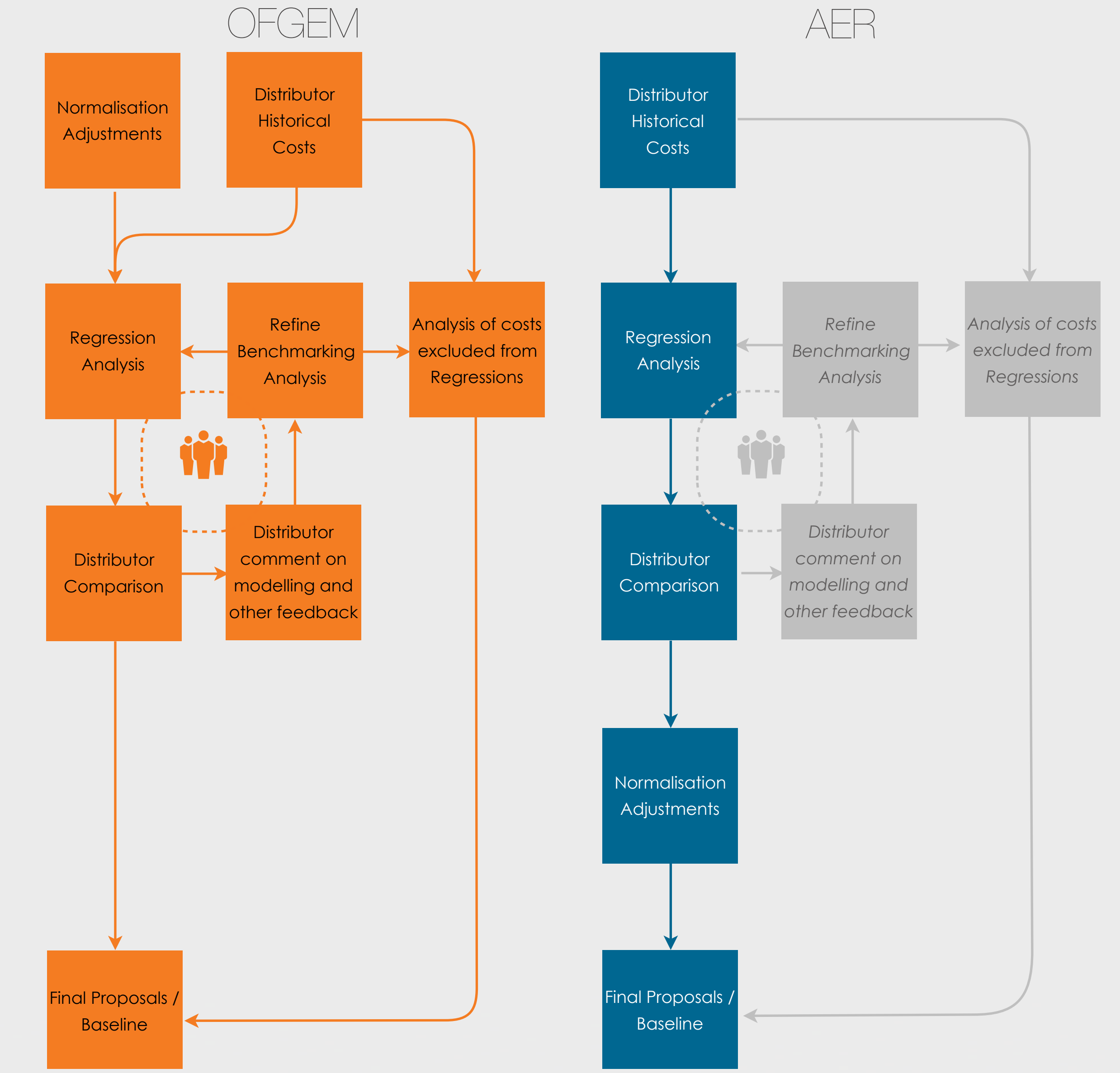
3.4.7 Explanation of inefficiencies

Our view is that the explanation of assumed inefficiencies has not been given appropriate consideration. The results of a limited number of models have been relied upon without appropriate investigation into the cause of the differences in efficiency scores. There has been little investigative analysis completed against the results to determine root causes of implied managerial efficiencies, other than references to labour productivity issues in government owned electricity networks.

3.5 Our view in summary

We consider that the weaknesses inherent in the MPFP and SFA benchmark measures, and the inability of the statistical practices to account for those weaknesses combine to such an extent that the outcome of the benchmarking is unreliable for the purpose of making a regulatory determination. Approaches such as that adopted by the AER take several years to evolve. When this is considered in conjunction with the observation that benchmarking models rarely (if ever) endure beyond one regulatory application in a consistent specification or use, we find that it is quite unreasonable to place reliance on the results of this benchmarking approach so early in its evolution. As a final comparison against better practice, the approach is considered alongside the OFGEM approach from the previous determination (DPCR5) in the graphic on the following page.

Exhibit 3.5: Operational cost methodology overview. *The OFGEM approach to determining "efficient" opex in their previous regulatory determination (DPCR5) compared to the approach adopted by the AER.*



Differences between the two approaches

Data normalisation:

- OFGEM has normalised the data before modelling the relationship between opex and cost drivers - this includes normalising data across the industry (for example IT & property costs have been removed) and also DNSP specific adjustments.
- Economic Insights has made the adjustments after econometric modelling has been undertaken.

Regression analysis:

- The OFGEM opex benchmarking relies on three different sets of results obtained by modelling expenditure from various levels of disaggregation - this means that operating expenditure is modelled at an aggregate level but also broken down into different expenditure groups (for example one group that is modelled includes opex for network policy, HR & Non-Operational Training, Finance & Regulation and CEO).
- Economic Insights has used a single aggregate expenditure model to determine the relative levels of efficiency between DNSPs.

DNSP consultation:

- OFGEM's opex benchmarking model involves extensive consultation with the DNSPs - "we have arrived at the final baselines through an iterative approach which has allowed extensive scrutiny by, and several rounds of interaction with, the DNOs and other stakeholders on the emerging results of our analysis".
- The AER have used the opex modelling results from their consultant without any feedback from distributors.

Why we consider the AER's benchmarking to be unfit for purpose

4

We believe that the AER has relied upon the results of the economic benchmarking analysis conducted by its consultant, Economic Insights, in determining adjustments to the base year expenditure of the NSW and ACT businesses. Our opinion is that such reliance is unsafe and premature, given the significant uncertainty in the results and infancy of the approach.

In this chapter

The following key points are relevant to this chapter on our views of the reasonableness of the approach

Huegin was asked to comment on the reasonableness of the AER's benchmarking approach as it relates to the NSW and ACT draft decision. We were instructed to address whether the AER has adequately identified and accommodated operating and environmental varaibles that may affect the results and any other factors that may contribute to the potential for error. Given the technical and procedural issues identified in the previous chapter, we focused on exploring the potential for error in the assumptions of inefficiency in the base year expenditure of the NSW and ACT businesses inherent in the modelling relied upon by the AER.

Key Point

There is undue confidence in what is a very narrow view of the range of results possible within the AER's own framework.

Sensitivity testing shows the variation in results possible through small changes in assumptions, however the AER has discarded all alternatives other than the narrow set it has relied upon. This practice is in contradiction to other regulators that combine models to account for bias.

Key Point

The lack of appropriate consideration of environmental variables causes bias against businesses operating under conditions not considered or tested by the AER's advisors.

Most notably, the lack of recognition of the physical asset differences, geographical differences and accounting policies of the businesses leads to erroneous conclusions of efficiency.

Key Point

Failure to consider the likelihood of multiple frontiers within the sample based on different classes and aggregation of network types invokes error in the measurement of efficiency gaps.

The use of an historical reference point for the frontier (the average over the data period) exaggerates the distance between current performance and the reference point for individual businesses.

4.1 The benchmarking models do not produce reliable estimates of efficiency

The AER have considered that the individual DNSP scores from their economic benchmarking models are a reasonable representation of relative efficiency. In our view the analysis that the AER has relied upon in recommending adjustments to NSW and ACT expenditure forecasts based on benchmarking is too limited to facilitate meaningful conclusions. Specifically, we consider:

1. There has been insufficient consideration of alternative methods and model specifications. The four models cited by the AER in the determination are each variants of a single model specification. Sensitivity testing reveals the full extent of the uncertainty in efficiency scores when comparing electricity network expenditure.

2. There is too much emphasis on a single, top-down benchmarking model. Alternative models that provide conflicting signals have been dismissed rather than recognised as evidence that any single model fails to appropriately describe the relative efficiency of the businesses.

3. The chosen model, its variables and the coefficients reflect a model that has poor explanatory power of the real operating costs of an electricity network.

4. There is a very real potential for statistical noise, measurement error and the influence of omitted variables to be interpreted as inefficiency. With such large variations in the efficiency scores from the model across the industry, the division between error, environmental differences and actual efficiency must be considered within the residual. Failure to properly distinguish between these components of the error term and other variables will lead to overestimation of the efficiency gap.

Each of the above points is discussed further in the following sections.

4.1.1 An insufficient number of benchmarking methods and model specifications have been considered

The number of model configurations possible for generating benchmarking results is vast given the choice of benchmarking methodologies, functional specifications and model specifications (definitions of each of these terms is in chapter 2). In section 3.4.6 of this report we referred to the Bauer consistency criteria as an often cited series of sensitivity tests that help put individual models into the context of the potential for biased results due to sensitivity to changing assumptions.

Section 2.3.1 outlined the concerns and qualifications with using international data. Nevertheless, to test the assumption that the benchmarking models do not produce consistently reliable results, we ran a number of alternative benchmarking methods and models using the data set used in the Economic Insights SFA model. The techniques used include opex MPFP, DEA and alternative SFA models and the results highlight the sensitivity of benchmarking results to changes in model specifications and assumptions. We have compared:

1. The results from the Pacific Economic Group COLS model conducted on behalf of the AER[17];

2. The results from the Economic Insights OPFP and SFA models relied upon by the AER in its decision; and

3. DEA models, alternative SFA and opex MPFP models that could have been analysed by the AER (DEA was not employed by the AER, and alternative model specifications were discarded).

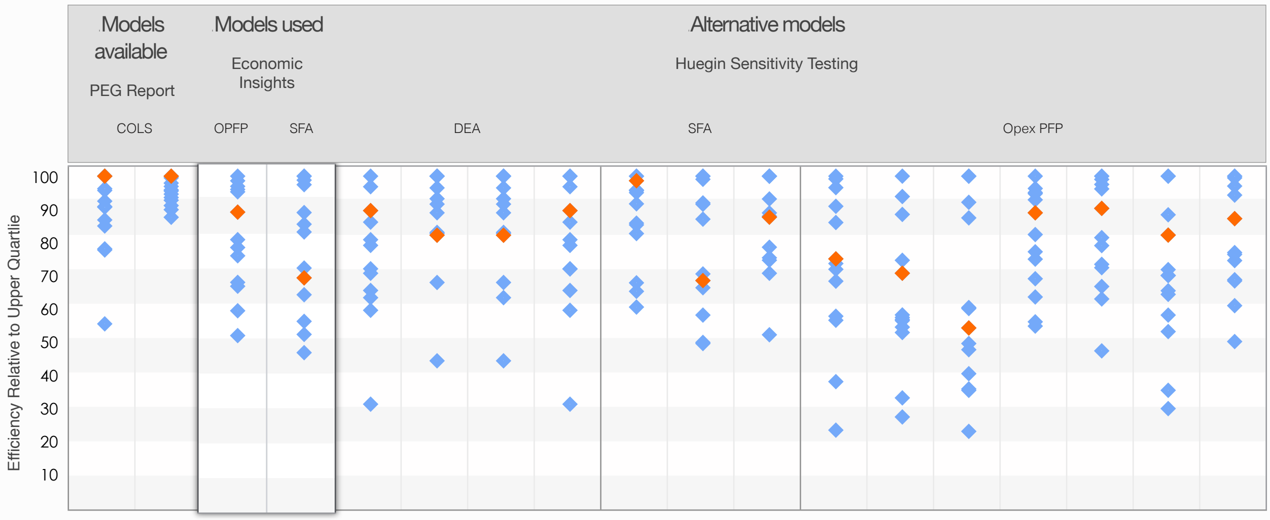
Annex C has details of the model specifications and techniques used for our tests. In summary, we used variations in the model input and output variable specification for the opex MPFP models, variations in the input and output specification and scale assumptions for the DEA models and for alternative SFA models we used the same specification as Economic Insights, but varied the error term assumption. The opex MPFP specifications were based on specifications considered by Economic Insights and the AER in previous studies (such as the Expenditure Forecast Assessment Guideline and the AER Draft Annual Benchmarking Report).

The Pacific Economic Group models are provided for context and contrast only. These models employed a different data set (Australian and U.S. networks), we have simply translated the results from the referenced report for comparison purposes. These models did not form part of our sensitivity testing, but do contribute to the argument that reliance on a narrow range of models and assumptions invokes undue risk.

Exhibit 4.1 presents the results of our analysis. Each column of dots represents the position of the 13 Australian businesses based on the efficiency score (normalised to 100%) for each of the 18 models analysed. Endeavour Energy's position is marked in orange, showing the significant variation in efficiency score across the model.

[17] Pacific Economics Group, "Database for Distribution Network Services in the US and Australia", 21 August 2014

Exhibit 4.1: Model sensitivity tests. *Australian electricity distribution network benchmarks of efficiency are particularly sensitive to modelling assumptions.*



The results in exhibit 4.1 can be placed in context of the Bauer mutual consistency criteria described in chapter 3. The results are shown in exhibit 4.2 on the following page. Exhibit 4.2 shows the variation across the models in:

1. **Efficiency scores**: The difference in the maximum and minimum efficiency scores for all businesses is between 40 and 68%. Every single business has at least one score of above 90% and at least one score of less than 60%.

2. **Efficiency rankings**: Most businesses register a ranking spread of at least seven positions across the various models.

3. **Top and bottom performers**: No business appears in the top four or bottom four performers for all 18 models. 11 of the 13 businesses appear in both the top and bottom four on at least one occasion.

In our view, this analysis demonstrates that the models fail the Bauer criteria. At the very least it raises adequate alarm that the results of any individual model should not be relied upon or used deterministically.

An observation of note within the analysis is that Ausgrid appear on the top of the rankings for two of the DEA models tested[1] - a technique identified in the Explanatory Statement for the Expenditure Forecast Assessment Guideline[2] as a useful cross-check of the productivity modelling, yet omitted from both the annual benchmarking report and the draft decision.

The AER's Explanatory Statement for the Expenditure Forecast Assessment Guideline recognised the imperative for robust sensitivity testing (see below). We consider that our analysis in this section demonstrates the level of sensitivity that the benchmark models are exposed to. We understand that the AER (through their consultants) has tested many model specifications to arrived at the preferred model. However we consider that the manner in which most models have been discarded in favour of the most recent model demonstrates selectivity, rather than acknowledgement of inherent sensitivity.

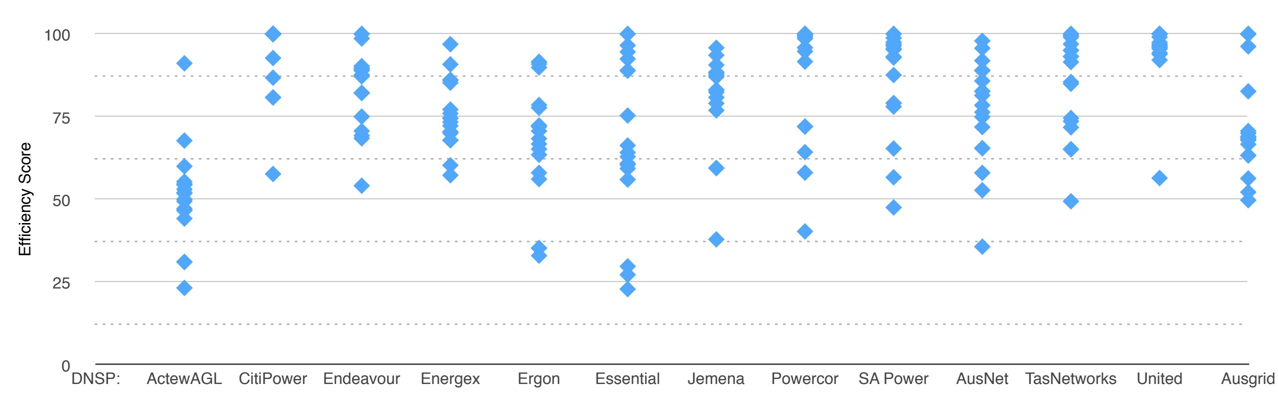
> ***...We will perform sensitivity analysis on model specifications, benchmarking methods, and changes in key assumptions to test the robustness of the results.***
>
> ***...We consider sensitivity analysis is a critical process in developing and finalising our model specifications.***
>
> ***...Sensitivity analysis is a method for testing a model to identify where there may be sources of uncertainty. It is an important step in testing the robustness of our economic benchmarking analysis.***
>
> ***...We will test multiple model specifications for each economic benchmarking technique.***
>
> - AER, Better Regulation, Explanatory Statement, pp163-165, "Expenditure Forecast Assessment Guideline", November 2013

[1] Both these DEA models use an assumption of Variable Returns to Scale

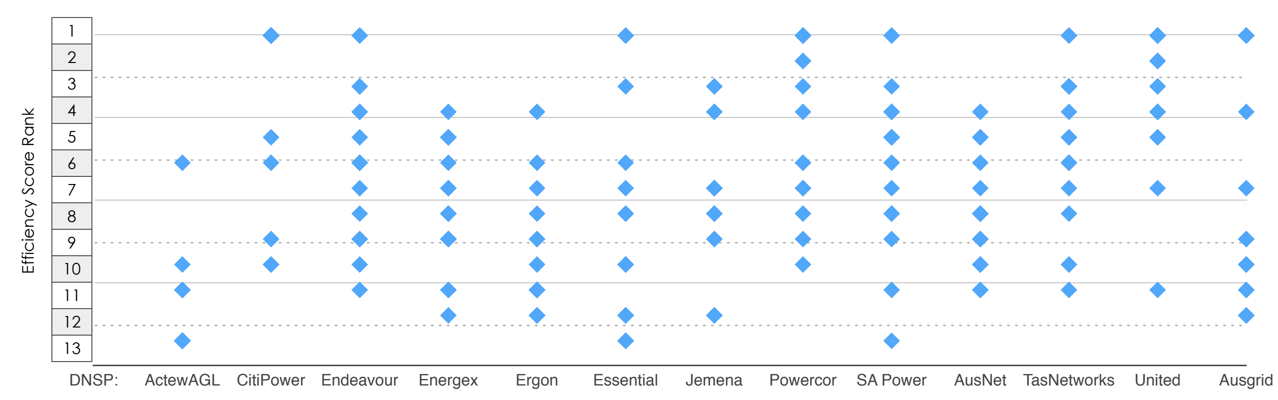
[2] AER, Better Regulation, Explanatory Statement, p124, "Expenditure Forecast Assessment Guideline", November 2013

Exhibit 4.2: Bauer consistency criteria. *The first three criteria of mutual consistency have been analysed. The 18 models tested for sensitivity do not produce consistent results, rankings or groupings.*

Bauer Level 1 - Do the efficiency scores stay the same across models?

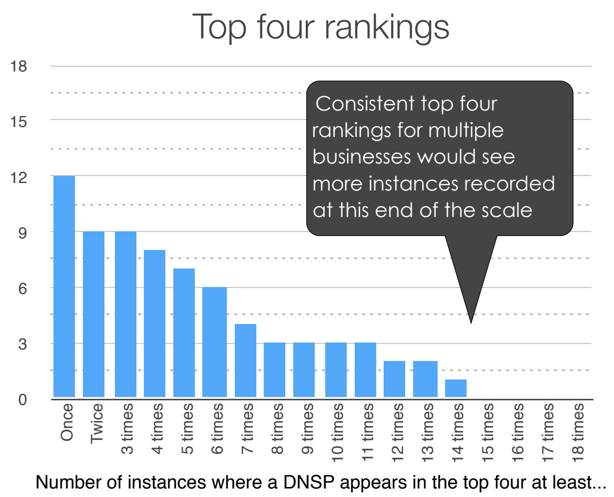


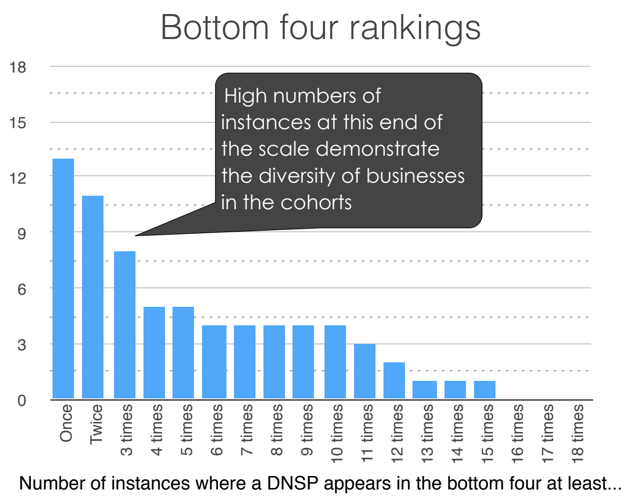
Bauer Level 2 - Do the efficiency rankings stay the same across models?



Bauer Level 3 - Do the models produce consistent groups of best and worst performers?

Cohorts of top and bottom four were selected for the 13 businesses to assess consistency of best and worst performers.





4.1.2 Too much emphasis on a single model

Huegin believes that the level of confidence the AER is placing in the econometric SFA model is unwarranted. Other regulators have avoided SFA due to the large dataset requirements. Those that do use it have much larger numbers of businesses to analyse and combine SFA with other methods[20].

> ***Given these potential issues, it would seem that the use of the SFA method to benchmark cost efficiency for regulatory applications should be approached with caution. Ideally, multiple periods of data on a large number of energy businesses in a sub-sector would be available for the analysis. In addition it may be prudent to follow the approach of academics and regulators that have applied the SFA benchmarking method. That is to undertake cross-checking of the SFA efficiency or productivity estimates (or rankings) against different model specifications and assumptions, and against different benchmarking methods such as DEA, OLS and its variants.***
>
> - ACCC/AER, p106, "Working Paper No.6", May 2012

Whilst the AER may consider that the introduction of international data facilitates the use of SFA, this practice introduces a raft of other problems with the modelling, including:

1. Reliance on Ontario and New Zealand data significantly restricts the level of consideration given to the impact of environmental variables on operating costs as the equivalent environmental variables are not available in those jurisdictions;
2. The potential for error and bias is amplified, due to the use of data which is collected under a different measurement and reporting framework with different accounting structures and regulatory requirements.

Every model will exhibit levels of bias and error when attempting to fit a simple cost function to a complex group of businesses operating in diverse conditions (this bias will favour Ontario and New Zealand given the disproportionate number of data points from those jurisdictions). This in itself is not a terminal flaw in the application of benchmarking. However without appropriate consideration of the statistical and environmental noise in a single, or limited number of models, the resultant inferences of relative efficiency are unreliable. As shown in the previous section, the potential range of outcomes with small changes in the model specifications or techniques is broad.

To address the effect of network heterogeneity and sensitivity of benchmarking results to both model specification and technique selection other regulators use geometric means, or some other form of weighting and combination of results from various model forms and methods to avoid bias in a single model. The table on the next page highlights the exhaustive analysis conducted by OFGEM during the DPCR5 expenditure review in 2009. OFGEM used a combination of both top-down and bottom-up benchmarking and a number of different cost drivers before arriving at relative efficiency scores. In particular, OFGEM used a number of models that accounted for different outlier businesses and also weighted the results of different models with the aggregate expenditure benchmarking results receiving a weighting of 9.09%[21].

The AER has not only placed disproportionate weighting on a single top-down model, but it has not taken into consideration other models available to it which cast significant doubt on the reliability of the results derived from its preferred model. This includes the modelling and results presented to it by another consultant, Pacific Economics Group[22]. Better regulatory practice dictates that an approach that balances the outcomes of a number of different models is appropriate, as it recognises that each model exhibits some level of bias. Therefore:

1. Disproportionate weight should not be placed on any single model; and
2. Where inconsistency in results exists, models should be combined in some way that at least mitigates the potential for bias in a single direction.

It is also worth noting the cost normalisation and exclusion of outliers in the OFGEM approach for DPCR5 shown in exhibit 4.3. Multiple groupings of costs are studied to remove the effects of unique, localised cost impacts and businesses considered outliers are excluded from some models to test sensitivity. A total of 40 models were used to inform OFGEM's conclusions. We note that in the current regulatory determination by OFGEM (RIIO-ED1), the approach has changed but the principle of combining multiple models has not. We also note in RIIO-ED1 that the maximum weight placed on any top-down economic benchmarking results is 25% and that the forecast of the network operator is also given a weighting of 25% in the final determination of costs[23].

---

[20] For example, regulators in Sweden, Finland and Germany have combined the results of SFA and DEA benchmarking

[21] Ofgem, "Electricity Distribution Price Control Review Final Proposals – Allowed Revenue – Cost Assessment Appendix", 7 December 2009, p92

[22] Pacific Economics Group, "Database for Distribution Network Services in the US and Australia", 21 August 2014

[23] OFGEM, "RIIO-ED1: Final determinations for the slow- track electricity distribution companies - Business plan expenditure assessment - Final decision", 28 November 2014

Exhibit 4.3 - Ofgem DPCR5: Weighting used for each set of analysis

Set of Analysis	Level of disaggregation	Difference to 'Core'	Free Weight for Driver	Outliers excluded	Weighting	Total
1	Top Down	Core			0.0120	
	Top Down	Core		Yes	0.0084	
2	Top Down	Driver - MEAV			0.0012	
	Top Down	Driver - MEAV		Yes	0.0008	
3	Top Down	Cost Base - including Non-load Capex			0.0006	
	Top Down	Cost Base - including Non-load Capex	Yes		0.0002	
4	Top Down	Cost Base - excluding Property			0.0120	
	Top Down	Cost Base - excluding Property	Yes		0.0048	
	Top Down	Cost Base - excluding Property		Yes	0.0084	
	Top Down	Cost Base - excluding Property	Yes	Yes	0.0034	
5	Top Down	Cost Base - excluding IT & Property			0.0120	
	Top Down	Cost Base - excluding IT & Property	Yes		0.0048	
	Top Down	Cost Base - excluding IT & Property		Yes	0.0084	
	Top Down	Cost Base - excluding IT & Property	Yes	Yes	0.0034	
6	Top Down	Cost Base - Regional adjustments LPN only			0.0006	
	Top Down	Cost Base - Regional adjustments LPN only	Yes		0.0002	
7	Top Down	Cost Base - excluding contractor indirects adjustments			0.0006	
	Top Down	Cost Base - excluding contractor indirects adjustments	Yes		0.0002	
	Top Down	Cost Base - excluding contractor indirects adjustments		Yes	0.0004	
	Top Down	Cost Base - excluding contractor indirects adjustments	Yes	Yes	0.0002	
8	Top Down	Cost Base - excluding Tree Cutting			0.0060	
	Top Down	Cost Base - excluding Tree Cutting	Yes		0.0024	**0.0909**
9	Single Group	Core			0.3788	
10	Single Group	Driver - Indirects - MEAV			0.0758	**0.4545**
11	Groups	Core			0.1043	
	Groups	Core	Yes		0.0417	
	Groups	Core		Yes	0.0730	
	Groups	Core	Yes	Yes	0.0292	
12	Groups	Driver - Group 1 - Load & Non-load costs			0.0104	
13	Groups	Driver - Group 1 - MEAV			0.0209	
14	Groups	Driver - Group 2 - Direct costs			0.0104	
	Groups	Driver - Group 2 - Direct costs		Yes	0.0073	
15	Groups	Driver - Group 2 - MEAV			0.0209	
	Groups	Driver - Group 2 - MEAV		Yes	0.0146	
16	Groups	Driver - Group 3 - MEAV			0.0521	
17	Groups	Driver - Underground Faults - Number of Faults			0.0104	
	Groups	Driver - Underground Faults - Number of Faults		Yes	0.0073	
18	Groups	Cost Base - Underground Faults - excluding Non-load Cables			0.0521	
19	Groups	Method - Group 3 on per DNO basis			0.0000	
	Groups	Method - Group 3 on per DNO basis	Yes		0.0000	**0.4545**

Notes:

1. MEAV = Modern Equivalent Asset Value; DNO = Distribution Network Operator; LPN= London Power Network
2. Table adapted from Ofgem, "Electricity Distribution Price Control Review Final Proposals – Allowed Revenue – Cost Assessment Appendix", 7 December 2009, p92

4.1.3 The chosen model is not reflective of industry costs

In the previous chapter we pointed to the fact that the model specification of input variables (customers, ratcheted peak demand and line length) is not strongly representative of the drivers of operating expenditure. The extent to which the model fails to explain the relationship between operating cost and network cost drivers is demonstrated through the consideration of the actual drivers of the various costs of an electricity distribution network. Exhibit 4.4 below shows the breakdown of operating costs of an electricity network and the associated cost drivers of that expenditure. A qualitative assessment of the relevance of the AER's adopted cost drivers against each category is also provided. This assessment highlights the potential for the SFA model relied upon to interpret legitimate changes in opex as inefficiency through:

1. Failure of the chosen model variables to explain variations in the costs due to the weak explanatory power of changes in those variables to the majority of network costs; and
2. Lack of consideration of the actual drivers of the expenditure through the model specification or environmental variables or other adjustments.

Quantitative analysis in section 4.5 of this report validates the view that failure to consider the appropriate drivers of expenditure presents the risk that top-down cost modelling will overestimate the potential for efficiency improvement.

Exhibit 4.4 - Cost categories, primary drivers and the AER's economic benchmarking model variables

Cost Category	Contribution to Industry Costs*	Activities	Primary Drivers	Customers	Peak Demand	Line Length
Maintenance Costs	17%	Inspection	Schedule, design, location	None	None	Moderate
		Routine corrective	Design, schedule	Moderate	None	High
		Non-routine corrective	Failure rates, design	None	None	Low
Emergency Maintenance	10%	Assisted	Exposure, proximity	Low	None	Moderate
		Unassisted	Weather	None	None	Moderate
Vegetation Management	13%	Audit	OH network, location, terrain	None	None	Low
		Clearance	OH network, location, vegetation growth rate	None	None	Low
		Tree trimming	OH network, location, vegetation growth rate	None	None	Low
Corporate Overheads	17%	Executive	Scale	Moderate	None	None
		Legal, HR, Finance	Employees, energy served, network service area	Moderate	None	Low
		Regulatory, insurance, debt and equity raising	Energy served, revenue	Low	Moderate	Low
Network Overheads	39%	Network control & systems operations	Location, complexity, level of automation	High	High	Moderate
		Network management	Design complexity, location	Moderate	Low	Moderate
		Network planning	Location, design complexity	Moderate	Moderate	Moderate

* This is the breakdown of opex cost categories of the aggregate industry expenditure. Data is sourced from the RINS, and we have used the RIN definition of each cost category - note that this includes the consideration of network control, planning and systems operations as "Network Overheads", costs that we would otherwise consider direct operating costs.

4.1.4 Alternative error term modelling demonstrates the volatility in the efficiency scores

In addition to model specification, econometric benchmarking is sensitive to the assumptions used to incorporate environmental variables. Whilst Economic Insights have incorporated the share of network underground directly into the cost function, another equally valid technique is to adjust the model error term (and therefore the measure of technical inefficiency) for the influence of the environmental variable. Whilst the choice of which method to use is a largely philosophical one, results suggest that whilst benchmarking rankings stay the same, estimates of inefficiency can vary significantly between the two techniques. One example is an analysis of the relative levels of efficiency between international airlines in which the author concludes:

> ***..we are comforted to find that the ranking of efficiencies do not vary greatly with the method selected but are concerned to find that the sizes of the estimated efficiencies do differ significantly.***
>
> - Coelli, Perelman, Romano, p271, "Journal of Productivity Analysis: Accounting for Environmental Influences in Stochastic Frontier Models: With Application to International Airlines", 1999

The theoretical differences between the two models were outlined by the authors as such:

- Case 1: Assume that environment conditions affect the shape of the production technology, or
- Case 2: Assume that environment conditions influence the firm's technical efficiency.

In principle, this approach (adjusting the efficiency scores for the effects of environmental variables) is similar to the two-stage approach suggested by the AER in their Economic Benchmarking Model Technical Report - "One way to model the impact of operating environment factors is to run two-stage regression analysis of raw MTFP results"[24]. Whilst we recognise that this adjustment is on the efficiency scores of MTFP benchmarking the principle is the same; modelling is conducted to obtain relative efficiency differences between DNSPs and then these scores are adjusted to take into account environmental differences between businesses.

Using this different assumption, and using the dataset used by Economic Insights, Huegin re-ran the SFA model with the environmental variable incorporated in the error term. The results, based on the raw efficiency scores were as follows:

- ActewAGL move from being 58% from the frontier firm to 40% from the frontier firm; when adjusted for inputs (according to the AER assumptions and process) and relative to the upper quartile ActewAGL would have received an opex reduction of 12% using this error term method;
- Ausgrid move from being 53% from the frontier firm to 27% from the frontier firm; when adjusted for inputs and relative to the upper quartile Ausgrid would have received an opex reduction of 9%;
- Endeavour Energy move from 38% from the frontier firm to 13% from the frontier; when adjusted for inputs Endeavour would have been within the upper quartile; and
- Essential Energy move from 42% from the frontier firm to 46% off the frontier firm; when adjusted for inputs and relative to the upper quartile Essential would have received an opex reduction of 33%.

Essential Energy are disadvantaged by the approach due to the large, rural nature of their network and the associated small amount of the network underground.

Given that there appears to be no definitive answer over which assumption is the "correct" one for accounting for environmental variables we believe the AER should recognise the significant variations in efficiency scores that can occur with the use of economic benchmarking and should place less reliance on the models for determining specific opex adjustments which are based on volatile estimates of efficiency.

[24] AER, Economic Benchmarking Model: Technical Report, June 2013

4.2 Lack of consideration of environmental variables leads to overestimation of efficiency gaps

The exhaustive data requirements of stochastic frontier analysis mean that the extensive data set collected by the AER, including a wide range of environmental variables, was largely ignored because the same information could not be collected for DNSPs in Ontario.

This limited dataset means that aside from differences in peak demand, customer connections, the share of a network underground and circuit length, the SFA model assumes the only other reason that opex varies between DNSPs in the NEM, New Zealand and Ontario is inefficiency. Whether or not this assumption is a valid one is irrelevant. The point is that the model used to determine the NSW and ACT businesses base year opex could not have been robustly tested and validated because the data is not available to conduct the required sensitivity tests. For example data limitations mean the following could not be considered as drivers of opex:

- asset age;
- climate factors;
- customer demographics;
- network design;
- network voltages;
- network accessibility;
- network utilisation;
- reliability standards;
- scale;
- policy and regulation; and
- the physical environment within which network operates.

Disregard for these drivers of network opex means that their effects on opex are aggregated and labelled as inefficiency, with some networks with favourable operating conditions appearing efficient whilst DNSPs in challenging conditions appear inefficient.

There is a limit on the number of variables that can be considered, so even with careful consideration of environmental variables, the residual will still exhibit a failure of the model to explain differences, rather than pure efficiency differences. Whilst the AER stress in the determinations for the NSW and ACT businesses that they have tested and validated their benchmarking techniques to ensure they are robust, it is difficult to believe that the effects of different model specifications have been tested given the only environmental data available for the three countries used in the analysis was the share of underground network.

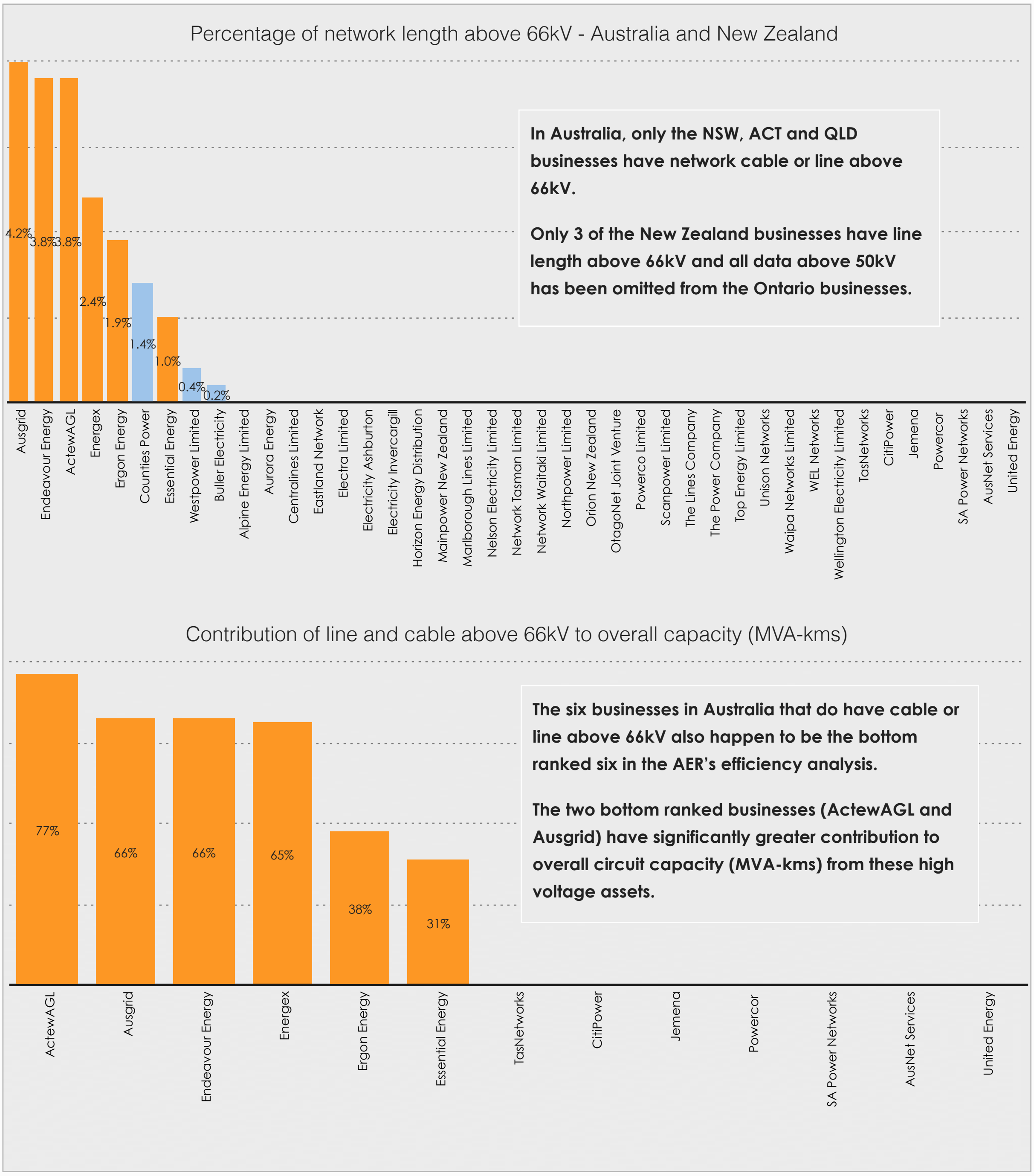
Four environmental variables that influence cost differences between networks - and that have a significant variation between frontier businesses and the NSW and ACT businesses - are discussed further in the following sections.

4.2.1 Capacity intensity

Huegin raised concerns with the MTFP model employed by the AER in the draft annual benchmarking report in August 2014 due to the bias against businesses with 132kV assets. Queensland, New South Wales and Australian Capital Territory businesses have 132kV assets in their networks due to the higher voltages at the point of transmission. South Australian, Victorian and Tasmanian businesses do not, nor do businesses in Ontario. In the MTFP model, where capital is represented by physical asset inputs, those businesses with 132kV assets were significantly penalised in the productivity score due to the method of measurement - MVA-kms of line. 132kV lines have orders of magnitude higher capacities than lower voltages, hence the input index for NSW, QLD and ACT businesses was artificially inflated. Economic Insights responded by splitting the overhead and underground line assets into below 33kV and 33kV and above. This does not mitigate the issue, rather it slightly diminishes the influence of 132kV assets. Victorian and South Australian networks do not have 132kV assets at all, so the assets and their associated opex should be removed from the inputs of the NSW, ACT and QLD businesses prior to modelling. We note that other regulators - including OFGEM and the Ontario Energy Board - remove costs associated with these very high voltage assets prior to benchmarking. We also note that the Commission for Energy Regulation in Ireland only benchmarks the single network under its regulatory control (which has 110kV assets) against other distribution networks that are also responsible for assets of 110kV and above.

Exhibit 4.5 shows that the presence of assets above 66kV in the network is almost unique to the NSW, ACT and QLD businesses and that the associated contribution to circuit capacity (measured as the MVA-km of overhead and underground) is closely aligned with the AER's productivity results.

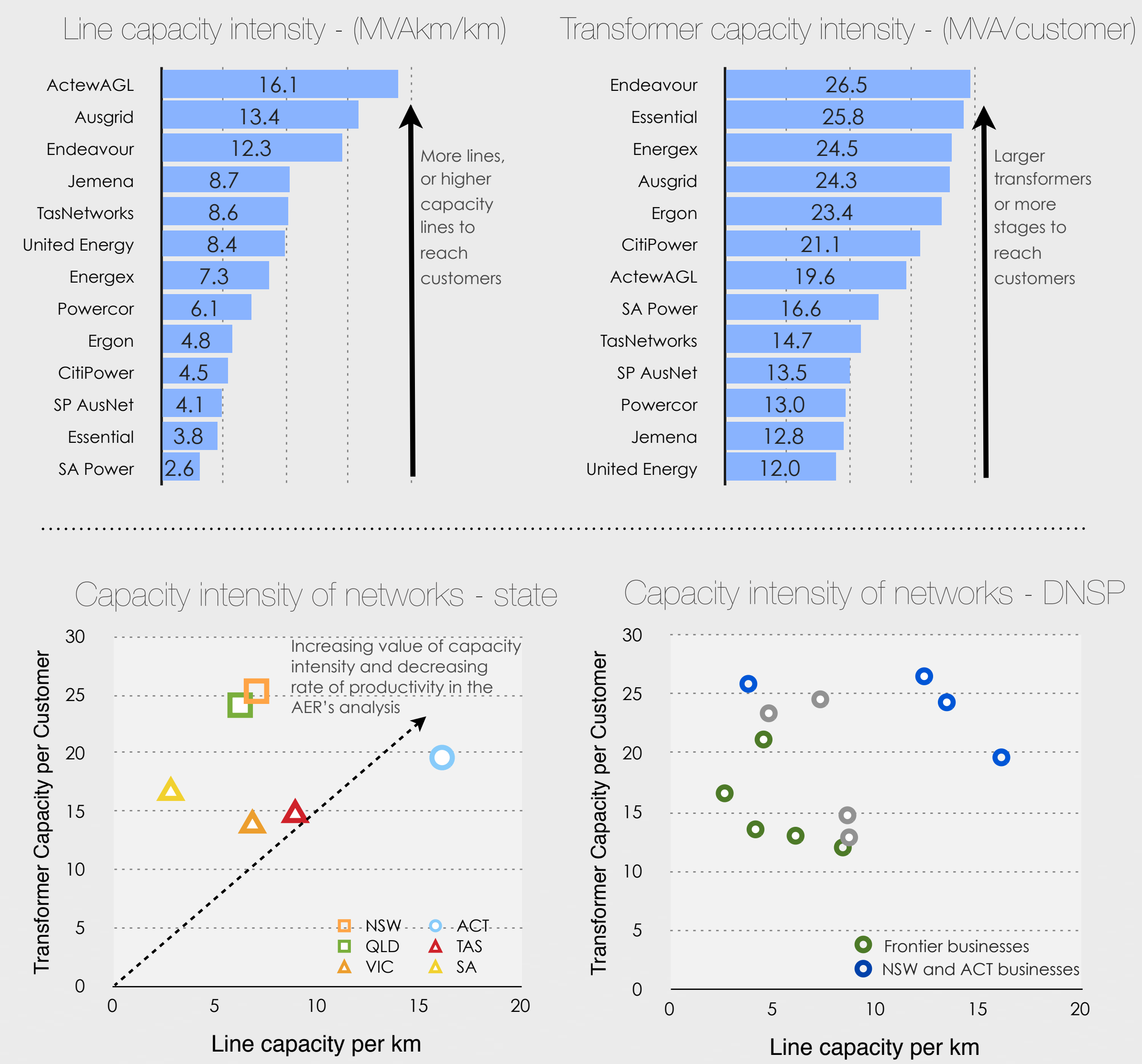
Exhibit 4.5: Network configuration - voltage and capacity. *Very few businesses have line or cable above 66kV, but for those that do it is significant in its contribution to capacity, if not length.*



Data sourced from 2013 Regulatory Information Notices (Aust) and NZ Commerce Commission

The particular issue of measurement of *physical* assets is less material to the benchmarking results since the AER has relied more upon the opex models (which exclude the physical asset inputs) in the draft decision, however the importance of network design should still be recognised as an environmental variable for opex. In the next section we discuss the impact of geographical distance. This, combined with higher voltage assets means that NSW and ACT businesses (and QLD) have much higher asset intense networks than their southern counterparts. Exhibit 4.6 shows that states that perform poorly in the AER's productivity analysis and DNSPs that perform poorly in the SFA Cobb-Douglas model have either a high transformer capacity per customer, high line capacity per km or both.

Exhibit 4.6: Capacity intensity. *Networks with high transformer capacities per customer and high line capacities per kilometre are those considered inefficient by the AER's models.*



The closer a data point is to the origin in the plots above, the less capacity intense its network is.

At the state level, the ACT is an outlier in terms of line capacity per kilometre (total MVA-kms of underground and overhead per kilometre total circuit length) and also rates poorly in the AER's productivity analysis. NSW and QLD transformer capacity per customer is highest amongst the states and they also perform poorly in the AER state productivity analysis. Victoria and South Australia have low line capacity per km and low transformer capacity per customer and perform best in the AER analysis.

At the DNSP level, there is strong correlation between the capacity intensity of a network and the productivity scores published by the AER according to the models that they have relied upon. Businesses with low transformer and line capacity per customer and km respectively are those on the frontier in the AER's model.

The existence of multiple stage transformation and high voltage line assets in the NSW, ACT and QLD businesses results from both the transmission system design in those jurisdictions compared to others and decades old planning and design decisions. They are not within control of the businesses to change to any significant degree.

4.2.2 Customer density and dwelling dispersion

The geographical and topographical features of the eastern seaboard of Australia are rarely considered in electricity network benchmarking. Customer density is often considered a proxy for differences in the geographical spread of a network's customers, however many networks with only twice as high a customer density above another have population densities tens to hundreds of times higher.

Customer density is often used as a normalisation factor between network businesses and the costs associated with operating in a variety of environments. The AER has considered that consideration of customer density can account for a range of factors that have a material influence on cost, including the number and exposure of assets, travel times, traffic management, asset complexity, proportion of overhead and underground and topographical conditions[25] . This is not the case at all. The AER's assumption is based on categorisation of the benefit or disadvantage to an urban or rural business only. Consideration must be given to:

1. The fact that customer density is an average over a network. Businesses such as Ausgrid and Endeavour have areas of very high density (Sydney), moderate density (Newcastle and Wollongong respectively) and very low density (the areas between Sydney and Newcastle and Wollongong respectively);

2. A highly meshed, small network can have the same customer density (measured as customers per km) as a longer, more radial network, however the travel distances to the assets will be much greater for the latter; and

3. Customer density (measured as customers per km) as a proxy for the factors listed by the AER does not account for the increase in depots and decentralised service functions required by networks servicing a larger area.

Customer density - the number of customers per linear kilometre of network - fails to account for the geographical spread of customers across a DNSP's service area. It is not coincidental that many productivity measures seem to indicate that Queensland productivity is lower than New South Wales productivity, which in turn is lower than Victorian productivity. The relative ratios of variance between these measures are similar to the differences in sparsity of the states. Just as customer density is often mistakenly assumed to be an appropriate normalisation factor for costs, sevice area is often overlooked. The AER has considered that service area is not a useful measure and that customer density is appropriate.

> **“**
>
> ***As the networks do not incur costs for areas that are un-serviced, customers per square kilometre of service area is not a useful measure for opex or service comparisons...***
>
> ***...As customer density per kilometre is a relatively easy concept to understand, we have adopted this as our standard approach...***
>
> ***...We are satisfied that an adjustment for customer density is not required.***
>
> - AER, p 7-106, “AER Draft Decision Endeavour Energy Distribution Determination, Attachment 07 - Opex”, November 2014

The relative ease with which customer density per kilometre is understood is not a reason to accept it as an appropriate approach for environmental variable measurement. And customer density per kilometre is not a useful proxy for identifying *distances between* customers, as claimed by the AER.

> **“**
>
> ***Customer density is a useful proxy for identifying the distance between customers.***
>
> - AER, p 7-104, “AER Draft Decision Endeavour Energy Distribution Determination, Attachment 07 - Opex”, November 2014

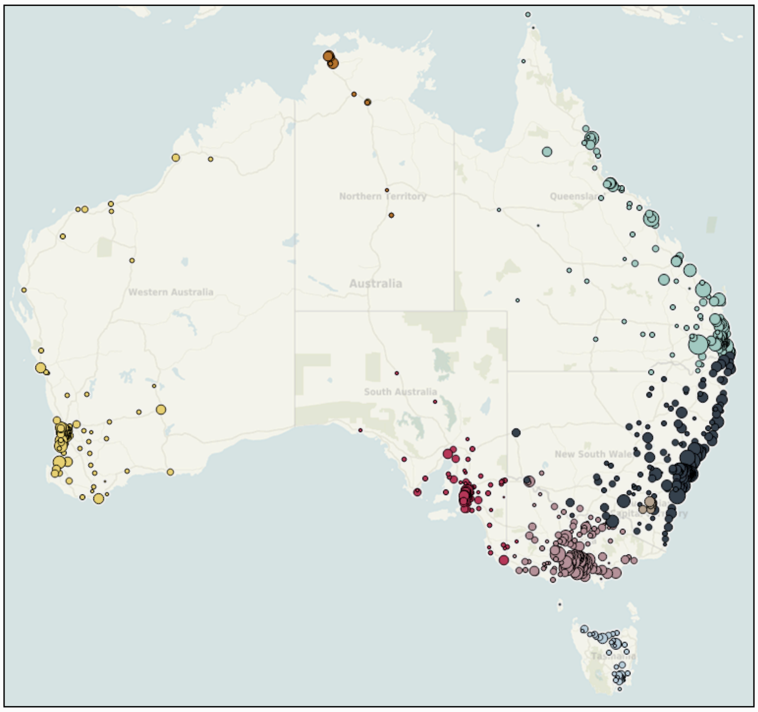
Customer density measured by line length does not account for distances between customers. A one hundred kilometre long feeder with 1,000 customers spread across it has the same customer density as a five hundred kilometre feeder with 5,000 customers at the end of it, however the costs to maintain and operate each would vary significantly.

Whilst it is true that large parts of rural networks are un-serviced, businesses are obliged to connect customers at the outer reaches of their network area regardless of the unpopulated areas between locations. To investigate the differences in the spread of customers across states, we studied the dispersion of dwellings across Australia - the number and location of dwellings by state. We did this by

[25] AER, Table A.18, p 7-106, “AER Draft Decision Endeavour Energy Distribution Determination, Attachment 07 - Opex”, November 2014

multiplying the number of dwellings in each location by the distance of that location from the state capital. This *is* a proxy measure for the distance between customers. Exhibit 4.7 illustrates the differences in dwelling density and associated productivity impacts across states and networks. The map in exhibit 4.7 shows that QLD and NSW both have a more significant spread of customers across the state area.

Exhibit 4.7: Dwelling dispersion. *The states with more of the population spread over greater distances fare poorly in the AER's productivity analysis. This is not inefficiency, rather an inability of the AER's models to explain the cost premium of transporting electricity over greater network areas.*

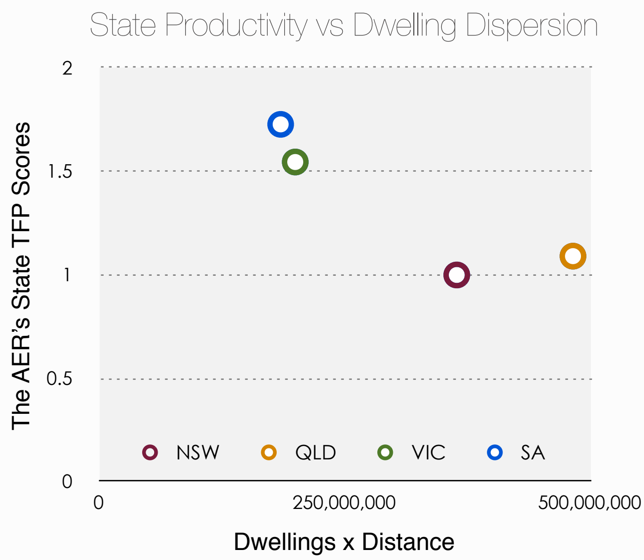


The graphic on the left shows the number of dwellings by location across Australia (the size of the bubble is the number of dwellings).

As shown, QLD and NSW have a significant number of dwellings spread over a vast distance. This not only increases operating and maintenance costs, it increases non-network and overhead costs.

A physical measure that highlights the variation across states is the number of dwellings times the distance from the capital. This is shown below.

State	Dwellings x Distance
QLD	481,617,473
NSW/ACT	363,283,504
VIC	198,832,789
SA	183,973,058



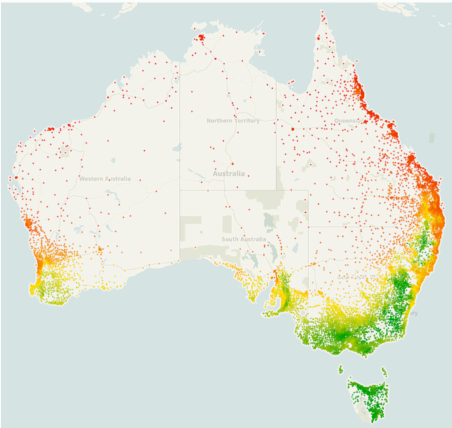
The graphic on the left shows the relationship between the mainland state TFP scores from the AER's analysis and the quantity and spread of dwellings.

The productivity analysis conducted by the AER fails to recognise that electricity is transported further across the state to more premises in NSW and QLD.

4.2.3 Climate and environment

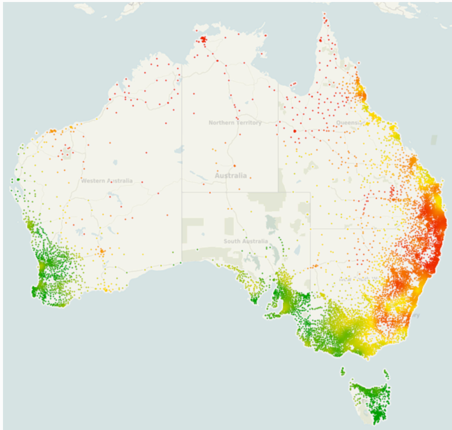
The AER nominated climate and environment factors as appropriate environmental variables against which to test productivity sensitivity in the Expenditure Forecast Assessment Guideline. Accordingly, the AER collected information on environmental variables in the RINs, but none have been used in the benchmarking analysis. This is in part due to the introduction of international data, but may also be due to the difficulty in distilling climatic differences down to a single variable across networks that have variable climates themselves. Nonetheless, it is noteworthy that these factors have not been considered in a jurisdiction with such extremes of climate and weather. Exhibit 4.8 demonstrates the materiality and variation in climatic factors across networks.

Exhibit 4.8: Climatic impacts increase to the north. *The large, expansive networks in NSW and QLD are more exposed to harsher weather climates than other states - there has been no consideration of this factor in the AER's analysis. The graphs below show important geographic and climatic differences across Australia. The colour of the map is the relative severity of the particular factor in that location - red being most severe, green most benign.*



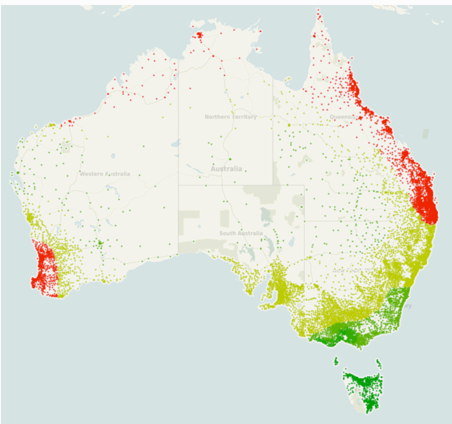
Average Temperature

Average temperatures are higher in Queensland and northern NSW.



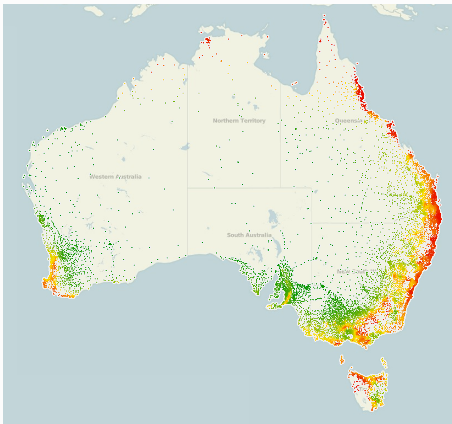
Severe Storms

NSW and south east Queensland have more instances of severe storms than other areas of the country.



Termites

Termite exposure is greatest in Queensland and northern NSW.



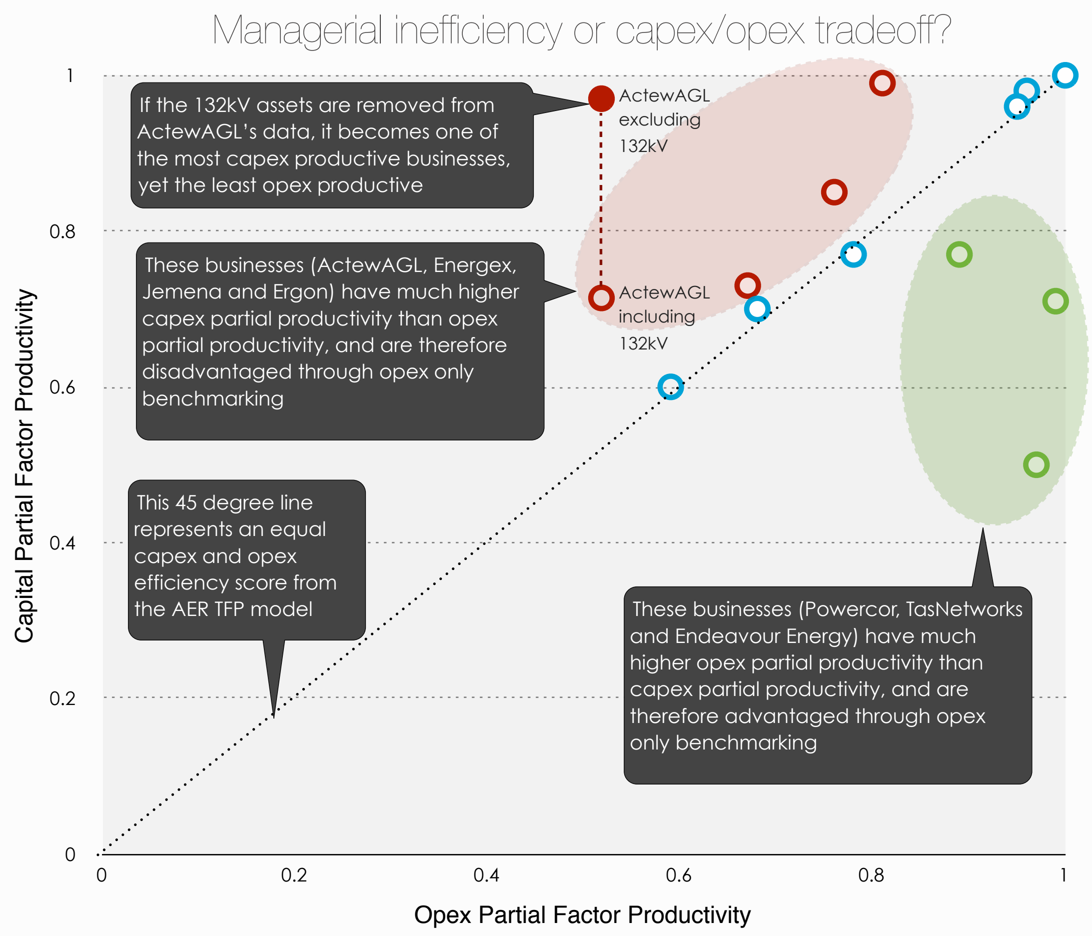
Rainfall

Rainfall is highest on the NSW north coast and south east Queensland.

4.2.4 Cost substitution

The AER has used benchmarking models to estimate the efficiency of operating expenditure. Whilst the AER have attempted to make adjustments to ActewAGL's expenditure, in particular, to account for capital and operating expenditure substitution, our view is that the models chosen overestimate managerial inefficiency and penalise businesses that have a more opex intensive network and business than capex intensive. One way we can see potential differences in cost structures (technology) is to compare the differences between opex and capital partial productivity factors obtained from the AER's benchmarking. DNSPs operating within similar technologies and cost structures should benchmark fairly similarly irrelevant of whether we use opex PFP or capex PFP as it would be counterintuitive for DNSP management to be efficient in one area but inefficient in another area. Exhibit 4.9 demonstrates the magnitude of the error in assuming that opex productivity scores are representative of managerial efficiency.

Exhibit 4.9: Managerial inefficiency overstated. *The comparison of opex and capex efficiency using the AER's chosen productivity models demonstrates the error in assuming the residual is inefficiency when studying opex only.*



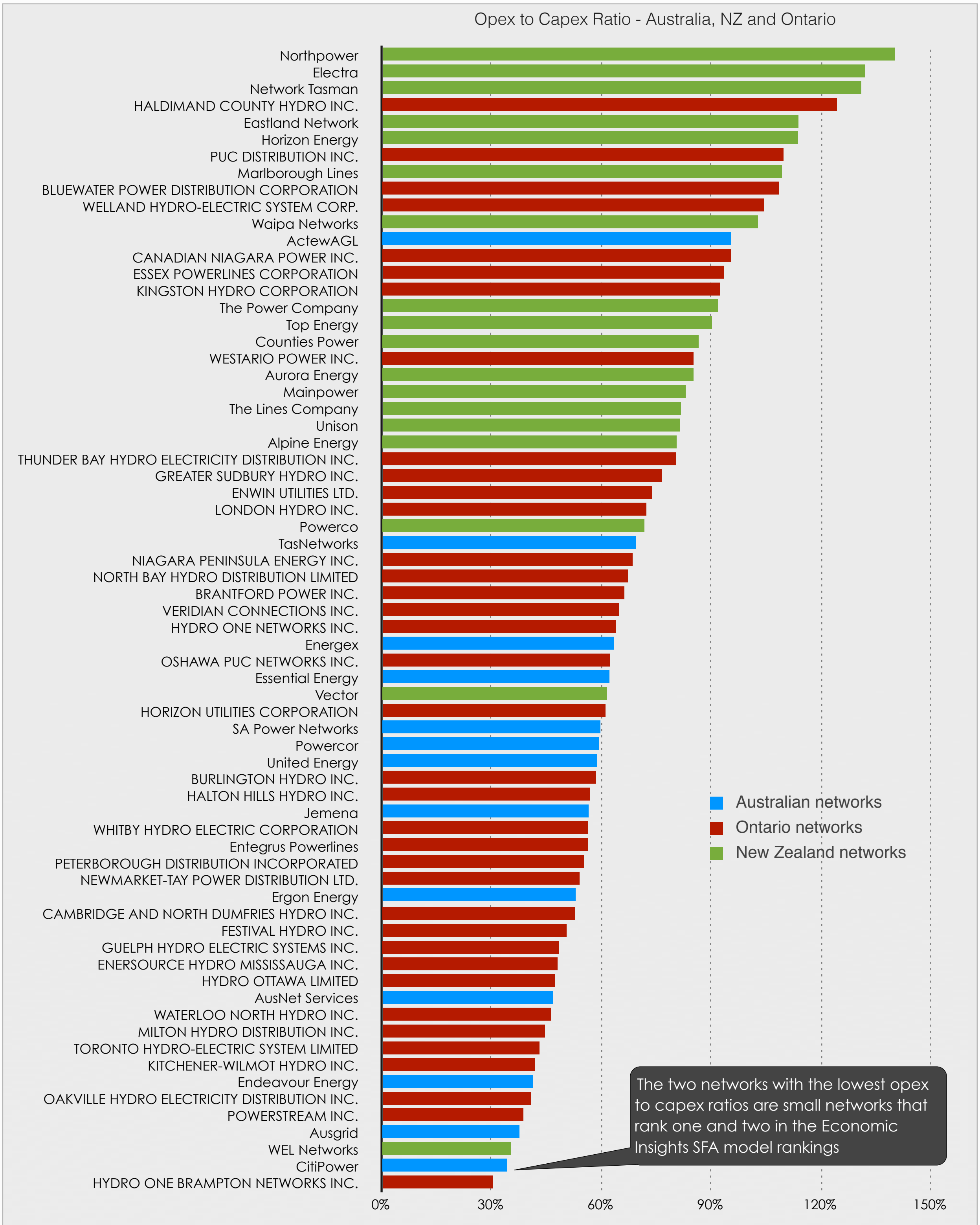
ActewAGL and Jemena (and Ergon Energy and Energex to a lesser extent) are particularly "penalised" by the AER's opex productivity model, due to their capex/opex allocation.

Endeavour Energy, Powercor and TasNetworks are favoured by the AER's approach due to their high opex productivity compared to capex productivity.

None of these observations are conclusions of actual efficiency, rather they demonstrate the error inherent in taking a partial model (opex or capex) without considering the other component of total costs. It is unlikely a business would be systematically inefficient in one cost category and efficient in the other - it is more likely an anomaly of the business allocation of cost between the two. ActewAGL are particularly affected by this issue; removing 132kV as a physical input (as many others don't have these assets) shows that ActewAGL are one of the most "efficient" businesses for capex, yet the least "efficient" for opex under the AER's model.

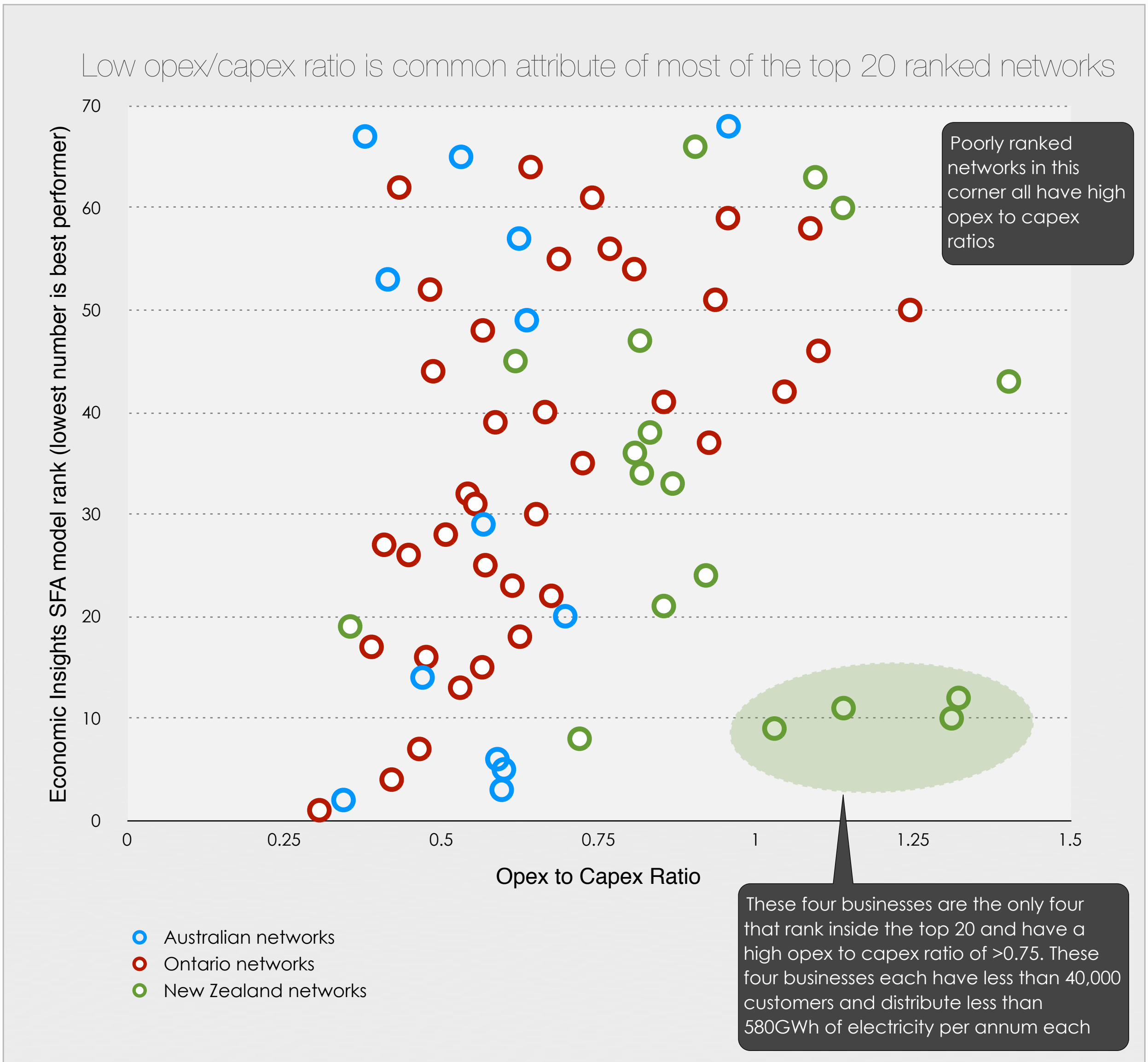
This issue of cost substitution is important when relying upon opex only benchmarking models, particularly if the businesses in the sample have different policies on capitalisation, or different technologies and/or assets that may be either more capital or operating expenditure intensive. Capex and opex tradeoffs also vary over time with investment cycles. Exhibit 4.10 shows the variation of opex to capex ratios for the networks of the NEM, Ontario and New Zealand. As shown, New Zealand businesses generally have high opex to capex ratios. Australian and Ontario businesses generally have lower opex to capex ratios (with the exception of ActewAGL).

Exhibit 4.10: Opex to capex ratio variation. *Networks with a high opex to capex ratio are advantaged by the AER's SFA model. Data is for the 68 businesses included in the Economic Insights SFA model. Data is sourced from regulatory reports for the most recent year.*



The relationship between opex to capex ratios and the rankings from the Economic Insights SFA model is presented at exhibit 4.11 below.

Exhibit 4.11: Opex to capex ratio and SFA efficiency rank. *With the exception of four very small NZ networks, high opex to capex ratios generally lead to worse performance in the model relied upon by the AER.*



The respective regulators in Ontario and Great Britain have both adopted total expenditure benchmarking in recognition of the cost substitution issue and consideration of the peril of relying on a single model. See OFGEM's statement on total expenditure benchmarking below.

> ***We intend to use totex benchmarking as well as more disaggregated benchmarking supported by technical/qualitative analysis as part of our toolkit for cost assessment. Totex models ensure that we consider DNOs" opex-capex trade-offs in our cost assessment. This means that we can identify the companies that have minimised total costs relative to specified cost drivers. Disaggregated models allow a less constrained and more intuitive specification of cost functions of different cost activities. We consider that using a variety of approaches acknowledges that there is no one correct model for assessing comparative efficiency but a number of plausible ones.***
>
> - OFGEM, p 9, "Strategy decisions for the RIIO-ED1 electricity distribution price control - Tools for cost assessment", 04 March 2013

4.3 The allowances that have been made for environmental variables are insufficient

Adjustments to the efficiency scores of the SFA model relied upon in the draft decision have been made to compensate for the inability of the Economic Insights model and limitation of the dataset to accommodate environmental variables. Whether the most material and appropriate environmental variables have been identified and considered is debatable. The AER state that they have considered 33 environmental variables and tested these for materiality.

> ***...we consider that we have captured all of the material variables to the extent that the economic benchmarking data and modelling permit. However, given they cannot account for every difference, we have also considered other operating environment factors that could potentially affect benchmarking comparisons in section A.5. Our analysis indicates that only a few of these factors appear to have a material effect on total opex and we have accounted for them appropriately in our draft decision.***
>
> - AER, p 7-45, "AER Draft Decision Endeavour Energy Distribution Determination, Attachment 07 - Opex", November 2014

For many of the variables, adequate data was not available to conduct such tests, particularly not for the international networks. Of those environmental factors that have been considered by the AER, they are underestimated and/or inadequately adjusted for. There is also no detailed analysis or explanation of the justification for deeming other variables insignificant in the draft decision to support the AER's claim that only a few of the factors have a material effect on total opex.

4.3.1 Expenditure impacts of environmental variables are underestimated

For all the consideration of environmental variables individually in the draft decision, the adjustment amount allowed for the collective influence of these variables is merely a subjective estimate.

> ***In addition, there are several other factors that we consider have little impact individually but, collectively, could potentially be more material. For the less significant operating environment factors, we are satisfied that it is appropriate to consider their impact on opex holistically...***
>
> ***...we are satisfied it is reasonable to incorporate a 10 per cent margin on input use into our adjustment for each service provider.***
>
> - AER, p 7-33 & 34, "AER Draft Decision Endeavour Energy Distribution Determination, Attachment 07 - Opex", November 2014

This approach undermines both the efforts to identify the influence of individual factor impacts and also the veracity of the econometric model results that form the starting point of the analysis.

Analysis of the environmental variables that were considered relevant by the AER also reveal a level of underestimation of their influence. The most material of these for Endeavour Energy is the subtransmission adjustment. The AER, as stated above, has incorporated a 10 per cent margin on input use. The amount that the AER has then adjusted for all environmental variables is $16.7 million (in 2013-14 dollars)[26].

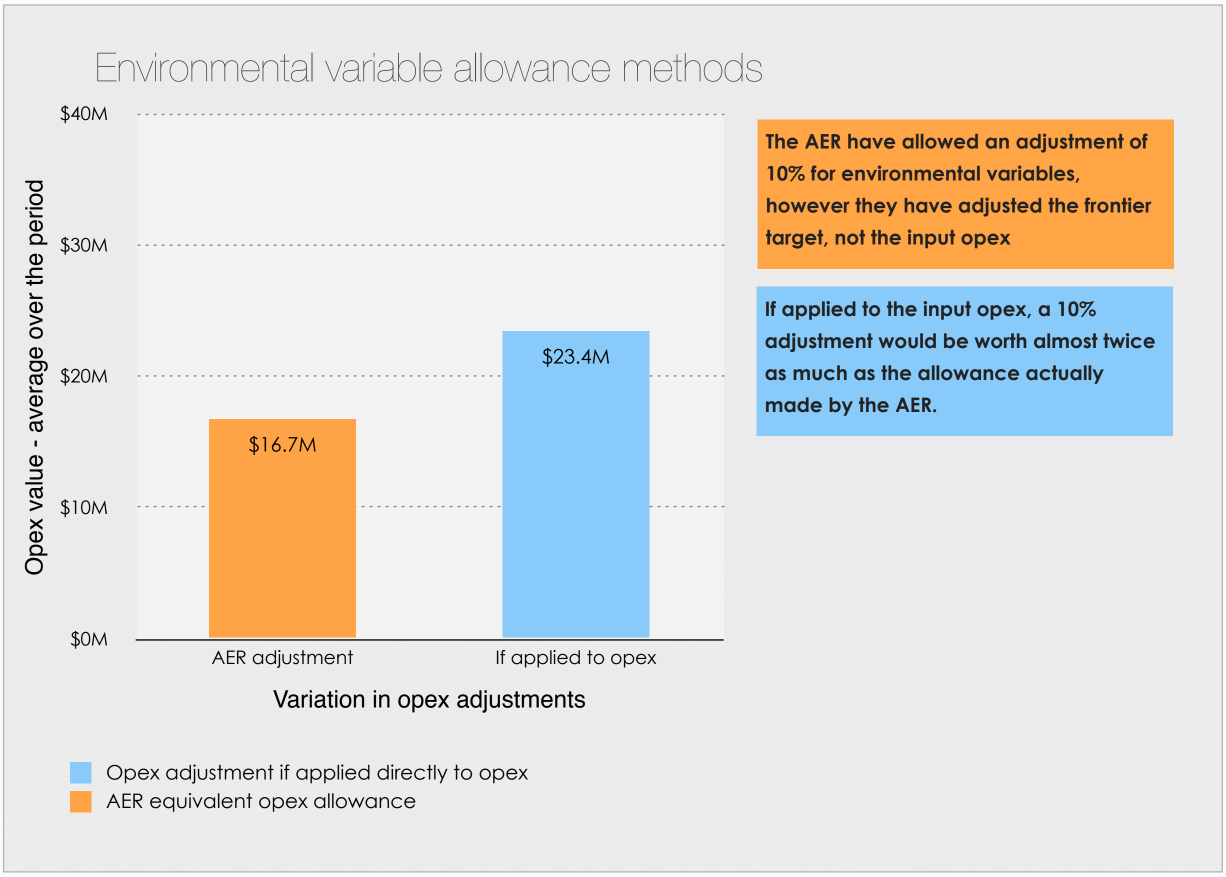
4.3.2 Environmental variables are inadequately adjusted for

One of the reasons that the adjustment for subtransmission assets is inadequate is that the AER has considered the split between 66 and 132kV assets and other lower voltages as the division of subtransmission and distribution networks. As discussed earlier in this document,

[26] AER, p 7-28, "AER Draft Decision Endeavour Energy Distribution Determination, Attachment 07 - Opex", November 2014

the NSW and ACT businesses (and QLD) have assets above 66kV, whereas none of the frontier businesses do. This disadvantages those businesses with these high cost assets in their network. However the other reason that *all* environmental variable adjustments are underestimated by the AER is the process of determining the adjustment amount. The AER's allowance for the margin on input use has actually been executed through an adjustment of the frontier, not the opex. Adjusting a number with no absolute meaning (the frontier efficiency score) by an arbitrary allowance (the environmental variable adjustment of 10%) is not an appropriate means of normalising opex factors. If the factors influence opex, opex should be adjusted. The difference between the allowance made by the AER using the method of adjusting the frontier position compared to the adjustment value had the opex been adjusted by 10% is shown in exhibit 4.12 below.

Exhibit 4.12: Opex adjustments - frontier vs input data. *The adjustments made by the AER for environmental variable impacts have been incorporated through an adjustment to the frontier target, not opex. The 10% allowance is equivalent to just $16.7 million.*



As stated earlier, the opex associated with the highest voltage assets in the NSW and ACT businesses alone accounts for more variation in opex compared to the frontier businesses (that do not have these assets) than has been allowed by the AER's adjustment for the collective 33 environmental variable. OFGEM excludes costs associated with high voltage assets prior to modelling. Of greater concern (given that the data is material to the results the AER has relied upon), the data from the Ontario businesses also excludes costs associated with assets over 50kV.

4.4 The frontier is not an accurate, nor appropriate reference point of efficiency

An efficient frontier is an academic concept. Not only does it rely upon the ability of the models to explain the majority of costs between businesses and over time, there are also other issues that must be considered, such as:

1. Are there businesses that are underspending and artificially shaping the frontier?
2. Is the timeframe for the analysis appropriate?
3. Are there signs of multiple frontiers in the data?

We believe that the analysis that the AER has relied upon is based upon a "false frontier" and explore the reasons for our opinion below.

4.4.1 The frontier employed is not reflective of industry costs

Reliance upon historical cost data coupled with the staggered regulatory cycle in Australia can cause progressive adjustments to operating expenditure over time to drive benchmark opex toward zero. There is no indicator available in the modelling to identify the point that the frontier pushes businesses below the threshold of sustainable operating costs. And just as there is no means of testing the true position of the efficient frontier, there is no means of testing whether the frontier used as a reference point in the AER's modelling represents an appropriate cost basis for the businesses under scrutiny. We have already explored issues such as the lack of opex explanatory power of the models, the absence of correction for environmental variables and the potential for error in the residuals. These all erode the confidence of inferences of efficiency using the model residuals. However the reference point - the frontier - also has a significant role in undermining the ability to make meaningful efficiency inferences.

Putting aside the fact that the residuals reflect the inability of the models to explain the differences between businesses, rather than efficiency gaps, this section explores the error inherent in the calculation of the frontier. Specifically, the analysis that the AER has relied upon:

1. Calculates a frontier from the econometric model based on the upper quartile scores and the average of the historical data (2006-2013);
2. Measures non-frontier businesses against this point;
3. Recognises that the base year expenditure has changed from the average, and therefore rolls the operating expenditure forward from the data mid point (3.5 years after 2006 and 3.5 years before 2013) for the NSW and ACT businesses; and
4. Makes adjustments for any loss or gain relative to the frontier for 2013 actual operating expenditure.

For Ausgrid and Endeavour Energy, who have reduced opex in real terms since the midpoint of the historical data period, this roll-forward mechanism provides them with a beneficial adjustment to the efficiency score the AER has relied on. For Essential Energy and ActewAGL, who have increased opex in real terms since the data midpoint, a penalty adjustment is made to the efficiency score that the AER has relied on.

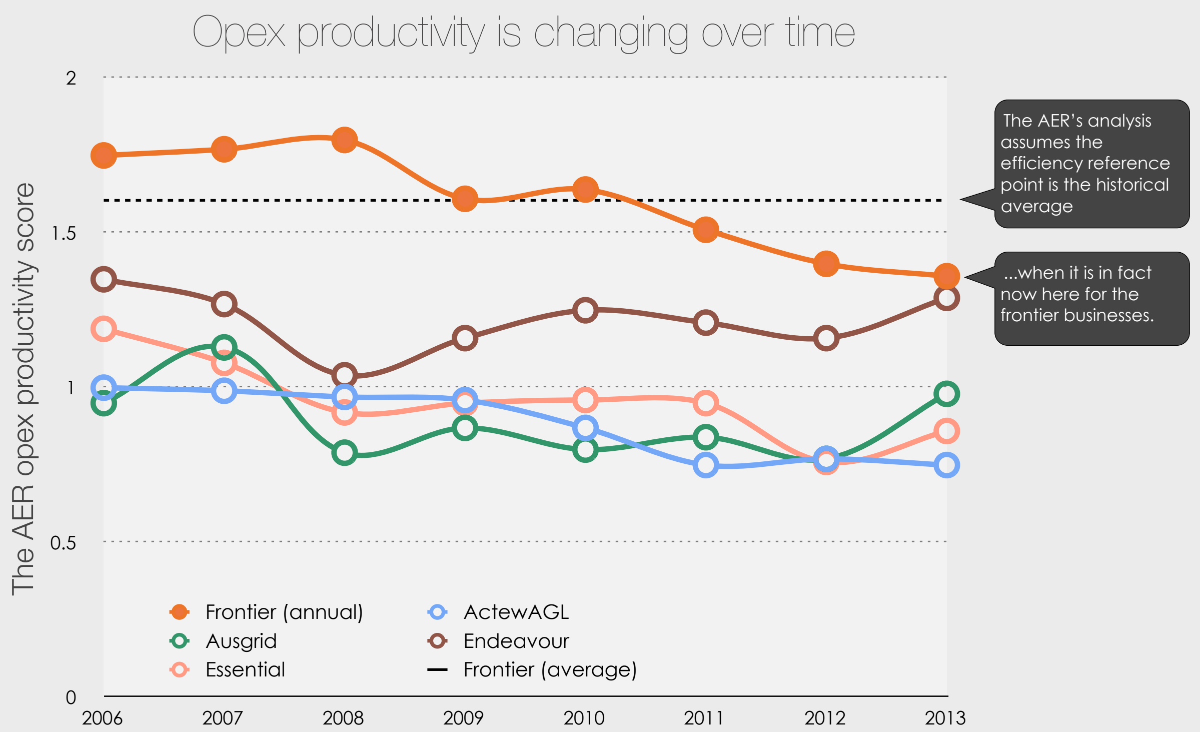
Whilst it may seem reasonable to roll forward the opex from the historical midpoint to more recent data for the businesses under scrutiny in the regulatory determination, failure to recognise that the costs of the businesses on the frontier under the AER's model have also changed has the effect of comparing recent expenditure of the NSW and ACT businesses against the frontier businesses 3.5 years previous. The consequences of this mismatch in comparison are compounded by the coincident nature of the midpoint of the historical data aligning with a point in time (halfway through 2009) at which the frontier businesses had not yet entered the current regulatory period. Analysis of the expenditure data of the frontier businesses shows that the operating costs of the frontier businesses have increased by 30% since the midpoint of the benchmark data period. The majority of this increase is related to increased maintenance and vegetation management expenditure in this regulatory period relative to the previous period. The graph in exhibit 4.13 shows the change in the opex partial productivity scores of the frontier firms since 2006, demonstrating the converging productivity of the frontier businesses[27].

It is worth noting that the primary driver of the convergence of productivity scores of the frontier relative to the NSW and ACT businesses is largely driven by increases in vegetation management expenditure by the frontier businesses (which in turn were precipitated by the Victorian bushfires). It illustrates the issue with omitted variable bias; the frontier appears to exhibit declining productivity, yet the profile is

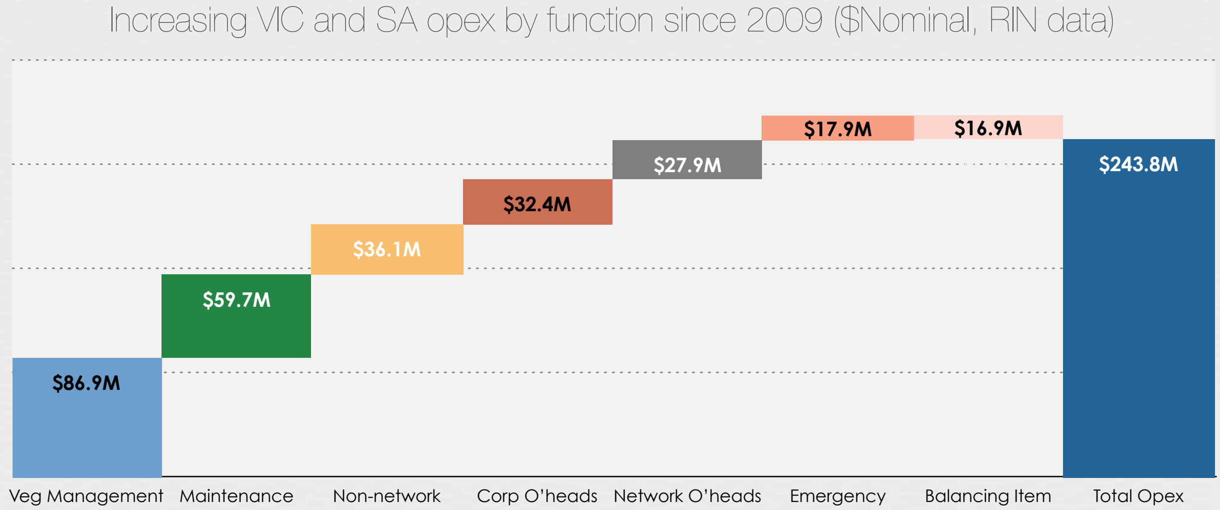
[27] The "frontier" score is taken to be the customer weighted average scores of each of the DNSPs identified as being on the frontier in Economic Insights *Economic Benchmarking of NSW and ACT DNSP Opex*

the outcome of the failure of the adopted model specification to explain increases in vegetation management - which is not driven by customer numbers, peak demand or line length[28].

Exhibit 4.13: Opex productivity over time. *The approach relied upon by the AER looks at productivity as an average over a historical period of seven years. It does not adequately recognise that the recent productivity results are much more converged than assumed.*



The changes in these graphs do not actually reflect efficiency change. They fail to recognise the actual drivers of opex as they reflect on the change in opex relative to changes in customer numbers, demand and line length. The graph does, however, demonstrate the issue with using long-term, historical averages of data for current efficiency measurement. The profiles above actually demonstrate the increase in vegetation management (the apparent decline in productivity is due primarily to the increase in opex, most of which has been vegetation management - see below) by the frontier businesses against the NSW and ACT businesses that have kept these costs relatively constant.

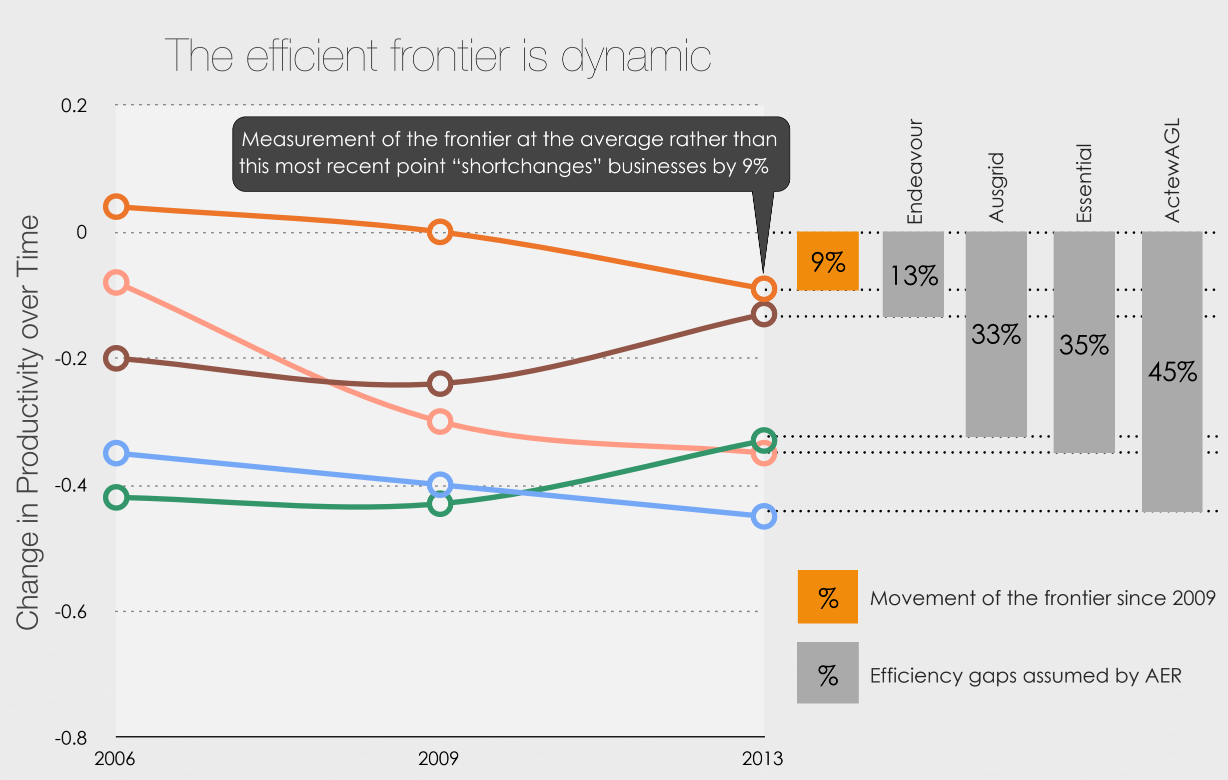


Data source: Category Analysis RINs for Victoria and NSW. Cost categories are as per the RIN guidance; the balancing item is the total value of the costs reported by the businesses in more than one category.

[28] Whilst there is some correlation between vegetation management costs and line length (as far as line length can approximate amount of overhead circuit), the changes in vegetation management costs over time have been primarily driven by fire related events during the period.

The declining productivity of the frontier businesses shows that there isn't a constant frontier over the period but a dynamic one driven by exogenous factors such as changing climate conditions, investment cycles and the staggered regulatory cycle. The graph in exhibit 4.14 highlights that the NSW and ACT DNSP's have moved relative to the average efficient frontier but so too have the businesses that themselves make up the frontier.

Exhibit 4.14: Frontier movement not considered. *The raw efficiency gaps assumed by the AER are based on current operating expenditure of the NSW and ACT businesses against a frontier that has moved since the reference point used in the analysis relied upon by the AER (mid 2009).*



Whilst the AER have allowed for the change in productivity of the businesses (which has a benefit for Endeavour Energy and Ausgrid, but a penalty for ActewAGL and Essential Energy) since the midpoint of their data, they have not recognised the current position of the frontier - 9% lower productivity than the midpoint. The efficiency gap assumed for all four businesses is therefore 9% more than it should be assumed to be. Even though Huegin maintains that the frontier is not reflective of the efficient level of expenditure (due to inappropriate cost model and lack of consideration of environmental variables), the comparison point should at least be aligned.

Had the AER compared the NSW and ACT businesses to its frontier at the most recent, common point in time, the adjustments it has made would be **9%** less than what they have assumed (due to movement of the frontier by 9% in 2013 relative to the historical average). This does not suggest that the frontier, even in 2013, is an appropriate reference point (due to the other issues raised in this report), however it demonstrates the significance of using averages over historical data to measure current efficiency- in particular, assuming a constant frontier over the period overestimates the distance between the NSW and ACT DNSPs in 2013 and the frontier in 2013.

4.4.2 A single frontier is unlikely in the Australian environment

The potential for unobserved heterogeneity (the presence of unmeasured environmental variables that influence the size of the residual) in the efficiency model residuals is high in the Australian environment, with its diverse operating conditions. With the inclusion of the international data, that potential is even greater. The influence of unobserved heterogeneity on the results relied upon by the AER is highly probable given the absence of appropriate consideration of environmental variables, weak explanatory power of the selected variables and inclusion of international data. Where such conditions exist and they cannot, or have not, been accounted for in the model itself (through environmental variable adjustments or data normalisation) it is good practice to recognise that multiple frontiers will exist amongst the modelled efficiency scores. This practice is known as latent class modelling, where businesses under observation in large datasets are classified into groups of similar attributes. Clustering algorithms can be used to determine appropriate groupings - these are well established practices that are also used for activities such as customer segmentation for marketing purposes.

Comparison of a business such as Essential Energy against a business such as CitiPower, where the only environmental variable considered in the model is the share of underground, presents compelling evidence that multiple reference points for classes of network is prudent. The presence of New Zealand and Canadian businesses in the data sample further enhances the case.

Examples of latent class modelling can be found in the benchmarking of U.S. transit systems, U.S. electricity transmission networks and European electricity networks.

In Norway, analysts found four classes of network amongst the 111 networks studied[29] - and Norway has far greater homogeneity amongst its networks and across its landscape than Australia. In Norway, it was found that conventional models underestimated cost efficiency and the latent class models had less risk of mistakenly capturing heterogeneity impacts as inefficiency.

> ***In general, the efficiency values are higher and more realistic than the corresponding scores of a conventional analysis performed in one step. The decomposition of the benchmarking process into two steps and the consideration of technology classes has reduced unobserved heterogeneity within classes and, hence, reduced the unexplained variance previously claimed as inefficiency. Therefore, conventional cross-sectional or pooled models might underestimate cost efficiency.***
>
> - Agrell, Farsi, Filippini & Koller, p 299 “Unobserved heterogeneous effects in the cost efficiency analysis of electricity distribution systems”, January 2013

Huegin conducted latent class modelling on the Australian, New Zealand and Ontario data used by the AER in its benchmarking. Using the large dataset of Australian, NZ and Canadian DNSPs (113 DNSPs over 8 years) we used latent class analysis to estimate the optimal number of technological groups the DNSPs operate within using circuit length, customer connections and share of underground as the DNSP characteristics.

The Akaike Information Criterion[30] was used to determine the optimal number of classes. We found the presence of four different technological groups within the dataset. These results suggest the presence of four distinct technological groups among the DNSPs in the dataset. Research has suggested that ignoring the presence of underlying technology differences results in overestimated inefficiency results as DNSPs are not benchmarked against businesses with comparable network characteristics.[31]

We note that latent class modelling overcomes the challenges of adjusting traditional benchmarking models specifically for environmental variables. This formed part of the conclusion of a latent class modelling study of U.S. transmission networks:

> ***We have also found that a simple latent class model is able to control for heterogeneity in firms’ operating environment without explicitly including environmental variables that regulators might find it very difficult or expensive to collect.***
>
> - Llorca, Orea & Pollit, p 15 “Using the latent class approach to cluster firms in benchmarking: An application to the US electricity transmission industry”, March 2014

[29] “Unobserved heterogeneous effects in the cost efficiency analysis of electricity distribution systems”, 2013. Agrell, Farsi, Filippini, Koller and “Efficiency Measurement using a latent class stochastic frontier analysis”, 2004 Orea, Kumbhakar

[30] The Akaike Information Criterion is a measure of the statistical quality of a model relative to other models.

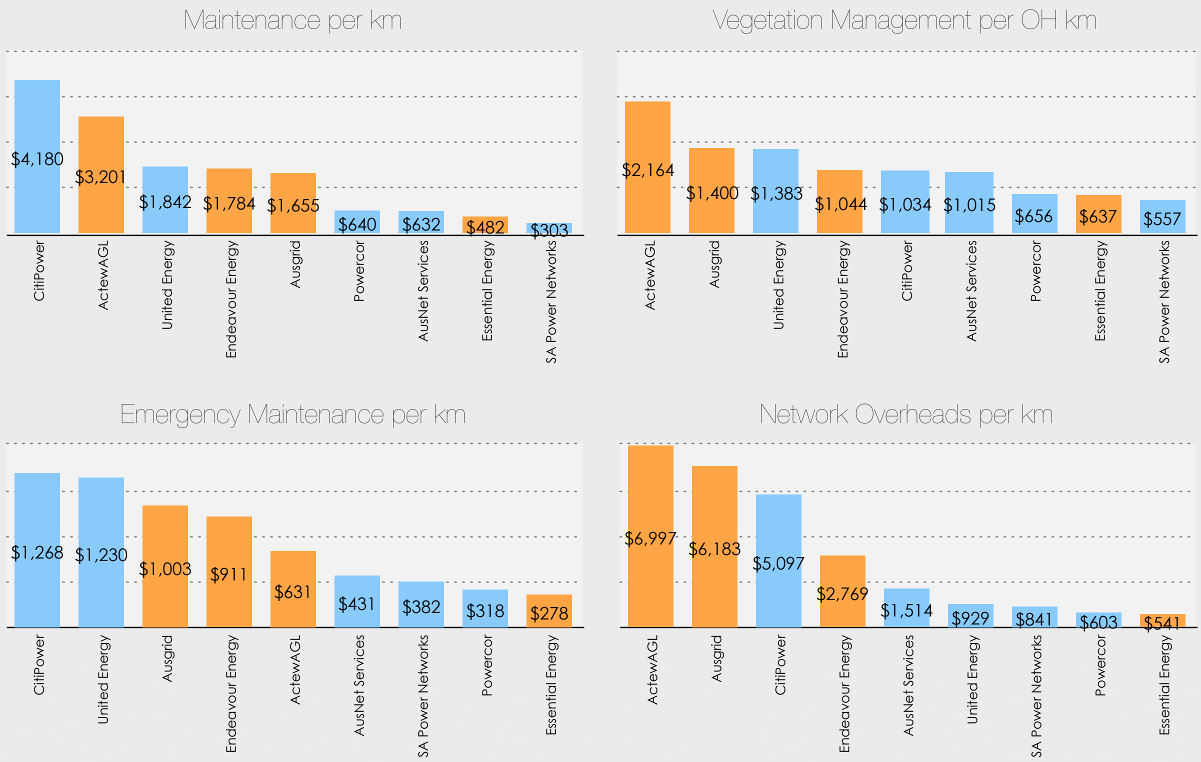
[31] “Unobserved heterogeneous effects in the cost efficiency analysis of electricity distribution systems”, 2013. Agrell, Farsi, Filippini, Koller and “Efficiency Measurement using a latent class stochastic frontier analysis”, 2004 Orea, Kumbhakar

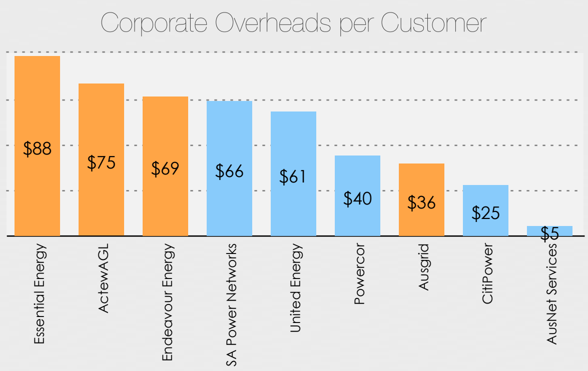
4.5 Bottom-up analysis demonstrates the contradiction with the AER's top-down SFA approach

The lack of engineering analysis or bottom up benchmarking compounds the exposure of the top-down analysis to bias. Without considering the performance from an engineering view, it is possible (and probable) that top-down results will underestimate total costs due to the lack of consideration of specific drivers of disaggregated costs. Partial productivity indicators can be misleading in isolation, particularly where there are differences in the cost allocation across categories between businesses. However, collectively, they are useful indicators of the presence or otherwise of systemic managerial inefficiency.

Exhibit 4.15 demonstrates that on a disaggregated cost level of partial productivity indicators, the NSW and ACT businesses do not systemically underperform against the frontier businesses. Partial productivity indicators are subject to the same sorts of errors of omitted variable bias and unobserved heterogeneity as economic modelling techniques. However they provide a useful counterpoint to the top-down analysis. At the very least these partial productivity indicators lead to the question of where expenditure reduction of the magnitude expected in the draft decision for NSW and ACT will come from.

Exhibit 4.15: Partial productivity analysis demonstrates the practical issue with meeting the AER adjustments.





When broken down into the primary cost categories, and using common denominators for partial productivity indicators, there is nothing to suggest that the NSW and ACT businesses are systemically overspending compared to the frontier businesses.

Many of the overhead costs reported by the NSW and ACT businesses are absorbed into the contract costs for direct maintenance activities for the frontier businesses, as the frontier businesses generally outsource more work.

Frontier businesses
NSW & ACT businesses

There is nothing to suggest in the balance of the partial productivity indicators that expenditure adjustments of the magnitude made by the AER to the NSW and ACT business's opex is warranted or possible. And this is prior to consideration of environmental variables such as network design differences, asset age, jurisdictional issues and differences in network service areas.

None of the frontier businesses are particularly good comparators for Endeavour Energy; United Energy would be the closest, although their network area is much smaller. Of note in exhibit 4.15, is that Endeavour Energy has similar partial productivity indicator performance to the other urban businesses, including CitiPower and United Energy in most categories. Corporate overheads is the only category where Endeavour Energy has a higher partial productivity cost than both CitiPower and United Energy, which is expected given the larger network area and the scope for the privatised Victorian businesses to allocate corporate costs across a larger group of businesses.

Summary and conclusions

Overall we believe that the level of error and uncertainty in the results of the benchmarking analysis relied upon by the AER in the draft decision for NSW and ACT prevents the models from producing valid estimates of efficiency.

Specifically we believe it is unsafe to rely upon the results, particularly deterministically, due to:

- The early stage of adoption and associated recent, numerous changes to the preferred model specification;
- The limitations and challenges inherent in the introduction of international data; and
- The demonstrable, material sensitivity of the results to minor changes in assumptions.

5

5.1 Our view is that the results are not a credible representation of efficient base year expenditure

We have demonstrated the inherent uncertainty in the AER's economic benchmarking approach throughout this report. In our view, the AER has placed undue reliance on the results and this has led to an underestimation of the efficient level of opex for the NSW and ACT networks.

5.1.1 We believe the AER has placed undue reliance on the benchmarking results

Our view is that the narrow consideration of model specifications, the limited alternative hypothesis and sensitivity testing and the coarse criteria for acceptance or rejection of a model (observations of scale or urbanity) cast sufficient doubt on the results of the benchmarking models to warrant extreme caution in their utility for predicting efficient opex of a DNSP. We find that the AER's reliance on the results in making the draft decision for the NSW and ACT businesses is contradictory to this view.

In our opinion, the AER's confidence in the results from the benchmarking analysis is disproportionate to the maturity of the approach in the context of a regulatory determination in Australia. The evolving model specification (and associated late changes), the reliance on historical data that is inconsistently reported and as yet insufficiently normalised and the introduction of international data all flag an approach that is indicative at best due to its infancy.

The staggered regulatory cycle presents the opportunity to accelerate the maturation of the process, models and data. However this will benefit the last jurisdictions in the cycle at the expense of the first - NSW and ACT. In our opinion, the results of the benchmarking and perhaps even the models themselves will be characterised by very different attributes at the end of the current regulatory cycle. This will, of course, be too late for the NSW and ACT businesses who will be faced with cutting operating expenditure to levels that may well be unachievable and certainly unsustainable.

5.1.2 We believe that the benchmarking results are not reasonably fit for the purpose of informing efficiency adjustments

Our view is that benchmarking of the type relied upon by the AER is useful for exploring the cause of differences in the efficiency score results of the models employed and to identify the physical and jurisdictional differences that drive those differences, efficiency or otherwise. Failure to acknowledge that the outputs of these models are subject to significant change with minor changes in the assumptions (and there is no consensus on the most reasonable set of assumptions) will lead to acceptance of the results as more meaningful than they actually are. We have highlighted the level of sensitivity around the models in this report, and it is further demonstrated through alternative models and results that have or will be presented to the AER prior to the final decision for NSW and ACT. But even if the current model specification and assumptions are accepted, there remains a significant level of doubt in the final results of the model due to second order treatments of the results and residuals and the data.

1. Potentially up to a **25%** error in the residual due to an alternative approach in modelling the error term (section 4.1.4);
2. An underestimation of the influence of environmental variables by several million dollars (exact amount depends upon which variables are considered (section 4.3.2);
3. At least a **9%** error in the adjustment assumptions based on a historical average frontier reference point (section 4.4.1) - this alone puts Endeavour Energy on the frontier in 2013;
4. The potential to underestimate cost efficiency of individual businesses due to the lack of consideration of latent classes and unobserved heterogeneity (section 4.4.2):

The consequences of the efficiency scores (and therefore the level of opex deemed efficient) being underestimated is an overestimation of the cuts to expenditure necessary to meet the AER's forecast. With the very real probability that the efficiency scores are underestimated, the decisions the businesses must make to achieve the allowable forecast of opex cannot be in the long term interest of consumers.

5.1.3 We believe that the deficiencies highlighted have not been appropriately addressed

In other jurisdictions, to overcome issues and deficiencies summarised above, regulators and benchmarking practitioners often:

1. Combine multiple model outcomes to mitigate bias of a single or narrow range of models;
2. Account for differences in scope and responsibility of functions between DNSPs by excluding certain costs before conducting economic modelling;
3. Use total expenditure benchmarking to overcome cost substitution issues and bias in capex/opex tradeoff circumstance;
4. Recognise the influence of unobserved heterogeniety, and therefore:
    a. Divide networks into clusters of similar, latent classes to avoid overestimation of inefficiency; or
    b. Adjust the econometric model error term to account for the influence of the environmental variable, rather than include it directly in the model cost function.

Our view is that the absence of similar efforts in the application of benchmarking that has informed the NSW and ACT draft decision seriously undermines the efficacy of the results.

The AER's apparent reliance on the benchmarking results to inform the efficient level of opex for the NSW and ACT networks, combined with the aforementioned failure to incorporate appropriate mitigation techniques to account for the inherent lack of stability of economic benchmarking models, renders the decision on the appropriate level of opex erroneous in our opinion.

> **"*The appropriate benchmark may also differ depending on the sensitivity of benchmarking results to technique and model specification. When there is uncertainty about the appropriate model specification and different specifications provide different results, it may be necessary to use the results cautiously.***
>
> - AER, Better Regulation, Explanatory Statement, p 128, "Expenditure Forecast Assessment Guideline", November 2013

Annex A - Letter of Instruction

A

5 January 2014

Jamie Blair
Huegin Consulting
Level 10, 2 Elizabeth Plaza,
North Sydney NSW 2060

Dear Jamie,

Letter of engagement – Networks NSW – AER Draft Determination

Ausgrid, Endeavour Energy and Essential Energy (referred to collectively as Networks NSW) are distribution network service providers in New South Wales, Australia regulated by the Australian Energy Regulator (AER) under the National Electricity Law (NEL) and National Electricity Rules (NER).

The AER made a draft determination of the revenue allowances for Networks NSW on 27 November 2014. This letter confirms your engagement in relation to Networks NSW's response to that draft determination and possible legal challenge of the final determination (which is expected in April 2015) (Response).

Scope of engagement

You are engaged by Networks NSW and ActewAGL, for the purposes of the Response, to:

a. provide economic analysis and advice;

b. prepare a written expert report (or reports);

c. appear as an expert witness for Networks NSW or ActewAGL (if required); and

d. undertake such other work as NNSW or ActewAGL may instruct you as the Response progresses.

A document outlining an list of questions that we require you to address in your expert report is set out in Attachment 1. These questions may be refined and developed, and added to, as the Response progresses.

A document outlining background on the regulatory regime relevant for the questions set out in Attachment 1 is included as Attachment 2.

Also enclosed is a copy of Practice Note CM7: Expert witnesses in proceedings in the Federal Court of Australia. Please ensure that your report complies with the requirements of Practice Note CM7, and also certify in your report that you have complied with Practice Note CM7.

Yours sincerely,

Catherine O'Neill
Group Manager – Strategy & Performance
Networks NSW

ATTACHMENT 1

LIST OF TOPICS REQUIRED TO BE ADDRESSED

The National Electricity Objective (**NEO**) set out in section 7 of the National Electricity Law is:

> *"The objective of this Law is to promote efficient investment in, and efficient operation and use of, electricity services for the long term interests of consumers of electricity with respect to—*
>
> *(a) price, quality, safety, reliability and security of supply of electricity; and*
>
> *(b) the reliability, safety and security of the national electricity system."*

Your report should address the following topics in the context of the NEO:

a. Comment on whether there is consensus on one right benchmarking method and what are some of the common challenges in carrying out such analyses

b. Review the benchmarking approach used by the AER. In addressing the foregoing, the consultant is asked to address whether the results can be relied on in setting regulatory allowance for operating expenditure

c. Comment on the reasonableness of the AER's benchmarking approach. In addressing the foregoing, the consultant is asked to address whether the AER has adequately identified, and reflected in the underlying analysis, operating and environmental variables that may affect benchmarking results, and any other factors that may contribute to the potential for error.

d. In relation to b above, comment on whether the benchmarking analysis should be used, and whether the deficiencies highlighted in b) and c) above can be addressed.

ATTACHMENT 2

BACKGROUND ON REGULATORY REGIME APPLYING TO ELECTRICITY DISTRIBUTION NETWORK SERVICE PROVIDERS IN NEW SOUTH WALES

INTRODUCTION

Networks New South Wales (**NNSW**) are the three *distribution network service providers* (**DNSPs**) in NSW – Ausgrid, Endeavour Energy and Essential Energy – regulated under the National Electricity Law (**NEL**) and Chapter 6 of the National Electricity Rules (**NER**). As such, NNSW were required to and did submit in May this year regulatory proposals to the Australian Energy Regulator (**AER**) for the determination of, among other things, their annual revenue requirements for the next regulatory control period (**Proposals**).

Chapter 6 of the NER sets out rules for the economic regulation of *direct control services* and *negotiated distribution services* provided by DNSPs. This regime requires the AER to determine the revenue allowed to be earned by NNSW for distribution services during each regulatory year, in accordance with the post-tax revenue model, described in Chapter 6 of the NER for each regulatory control period. In addition, a negotiating framework and negotiated distribution service criteria must also be determined by the AER. The process for making a distribution determination is set out in Part E of Chapter 6 of the NER.

DISTRIBUTION DETERMINATIONS

a. Under the NER, DNSPs must provide *direct control services* (which can be divided into *standard control services* and *alternative control services*) and *negotiated distribution services* on terms and conditions of access as determined under Chapters 4, 5, 6 and 7 of the NER (clause 6.1.3 of the NER). Relevantly, chapter 6 of the NER regulates:

b. for standard control services, the annual revenue requirements NNSW may earn for the provision of standard control services for which the AER must make a revenue determination (clause 6.3.2 of the NER); and

c. for negotiated distribution services, the requirements that are to be complied with in respect of the preparation, replacement, application or operation of NNSW's negotiating frameworks and the Negotiated Distribution Service Criteria (clauses 6.7.3 and 6.7.4 of the NER).

d. The making of a distribution determination is an economic regulatory function of the AER. As an economic regulatory function, section 16(1) of the NEL requires the AER to perform or exercise its function "in a manner that will or is likely to contribute to the achievement of the national electricity objective" set out in section 7 of the NEL being:

> *"The objective of this Law is to promote efficient investment in, and efficient operation and use of, electricity services for the long term interests of consumers of electricity with respect to—*
>
> *(a) price, quality, safety, reliability and security of supply of electricity; and*
>
> *(b) the reliability, safety and security of the national electricity system."*

e. In addition, if there are two or more possible decisions that will or are likely to contribute to the achievement of the national electricity objective, section 16(1)(d) of the NEL requires the AER to make a decision that it is satisfied will, or is likely to, contribute to the achievement of the national electricity objective to the greatest degree.

f. In addition, when making a distribution determination, the AER must also take into account the revenue and pricing principles set out in section 7A of the NEL:

"(2) A regulated network service provider should be provided with a reasonable opportunity to recover at least the efficient costs the operator incurs in—

(a) providing direct control network services; and

(b) complying with a regulatory obligation or requirement or making a regulatory payment.

(3) A regulated network service provider should be provided with effective incentives in order to promote economic efficiency with respect to direct control network services the operator provides. The economic efficiency that should be promoted includes—

(a) efficient investment in a distribution system or transmission system with which the operator provides direct control network services; and

(b) the efficient provision of electricity network services; and

(c) the efficient use of the distribution system or transmission system with which the operator provides direct control network services.

…

(6) Regard should be had to the economic costs and risks of the potential for under and over investment by a regulated network service provider in, as the case requires, a distribution system or transmission system with which the operator provides direct control network services.

(7) Regard should be had to the economic costs and risks of the potential for under and over utilisation of a distribution system or transmission system with which a regulated network service provider provides direct control network services."

B. OPERATING EXPENDITURE

The AER must determine whether it is satisfied that the forecast of required operating expenditure proposed by a distribution network service reasonably reflects the following criteria (clause 6.5.6(c) of the NER referred to as the **operating expenditure criteria**):

"(1) the efficient costs of achieving the operating expenditure objectives; and

(2) the costs that a prudent operator would require to achieve the operating expenditure objectives; and

(3) a realistic expectation of the demand forecast and cost inputs required to achieve the operating expenditure objectives."

The **operating expenditure objectives** referred to in clause 6.5.6(c)(1) of the NER above are set out in clause 6.5.6(a) of the NER as follows:

"(1) meet or manage the expected demand for standard control services over that period;

(2) comply with all applicable regulatory obligations or requirements associated with the provision of standard control services;

(3) to the extent that there is no applicable regulatory obligation or requirement in relation to:

(i) the quality, reliability or security of supply of standard control services; or

(ii) the reliability or security of the distribution system through the supply of standard control services,

to the relevant extent:

(iii) maintain the quality, reliability and security of supply of standard control services; and

(iv) maintain the reliability and security of the distribution system through the supply of standard control services; and

(4) maintain the safety of the distribution system through the supply of standard control services."

In deciding whether or not it is satisfied that the forecast of required operating expenditure proposed by a distribution network service reasonably reflects the following criteria in clause 6.5.6(c) of the NER, the AER must have regard to the following factors (clause 6.5.6(e) of the NER, referred to as **operating expenditure factors**):

"…

(4) the most recent annual benchmarking report that has been published under rule 6.27 and the benchmark operating expenditure that would be incurred by an efficient Distribution Network Service Provider over the relevant regulatory control period;

(5) the actual and expected operating expenditure of the Distribution Network Service Provider during any preceding regulatory control periods;

(5A) the extent to which the operating expenditure forecast includes expenditure to address the concerns of electricity consumers as identified by the Distribution Network Service Provider in the course of its engagement with electricity consumers;

(6) the relative prices of operating and capital inputs;

(7) the substitution possibilities between operating and capital expenditure;

(8) whether the operating expenditure forecast is consistent with any incentive scheme or schemes that apply to the Distribution Network Service Provider under clauses 6.5.8 or 6.6.2 to 6.6.4;

(9) the extent the operating expenditure forecast is referable to arrangements with a person other than the Distribution Network Service Provider that, in the opinion of the AER, do not reflect arm's length terms;

(9A) whether the operating expenditure forecast includes an amount relating to a project that should more appropriately be included as a contingent project under clause 6.6A.1(b);

(10) the extent the Distribution Network Service Provider has considered, and made provision for, efficient and prudent non-network alternatives; and

(11) any relevant final project assessment report (as defined in clause 5.10.2) published under clause 5.17.4(o), (p) or (s);

(12) any other factor the AER considers relevant and which the AER has notified the Distribution Network Service Provider in writing, prior to the submission of its revised regulatory proposal under clause 6.10.3, is an operating expenditure factor."

The AER's draft decision substantially reduces the forecast operating expenditure of the NNSW businesses when compared with the revenue proposals and historical operating expenditure allowances of those businesses. This is in part based on the outcomes of the annual benchmarking report, which determined that the NNSW businesses were not as efficient as other distribution network service providers and is without any transition to enable the NNSW businesses to adjust their practices to satisfy the reduced forecast operating expenditure allowance.

Annex B - Team member CVs

B

Jamie **Blair,** B.Eng (Chem)

Role: Project Lead

Jamie is a Director of Huegin Consulting Group and our project lead and electricity benchmarking expert. Jamie has significant experience in cost analysis and benchmarking in the electricity industry and often presents Huegin's work at industry conferences and academic forums. Jamie has extensive asset management experience, both in industry and consulting.

Relevant Skills

	Industry benchmarking
	Performance assessment
	Regulatory supporting, including revenue proposal analysis and review
	Risk management
	Safety reporting
	Maintenance and cost modelling
	Analytical decision support and statistical analysis

Relevant Experience

Jamie is an experienced engineer and consultant with specific expertise in the areas of investment analysis, cost analysis and performance benchmarking. His work is primarily for clients who own, manage or operate large physical assets. Relevant experience includes:

- Led over twenty independent benchmarking studies of domestic and international electricity networks.
- Facilitated the corporate strategic planning of an electricity distribution business and a utilities maintenance organisation.
- Developed the asset management frameworks for a major transport infrastructure manager and a large Defence weapons logistics management organisation.
- Led the analytical review of five recent regulatory determinations on behalf of network service providers.
- Developed and implemented the investment decision support framework and systems of a large network operator.
- Developed and implement the investment decision support framework and systems of a ports operator.

Professional Summary

Jamie Blair is a Director of Huegin Consulting. Jamie has 20 years of management and consulting experience across a number of industries including utilities, construction, military aviation, banking and finance and fast moving consumer goods.

Prior to joining Huegin in 2008, Jamie has worked in industry specialist consultancies, management consultancies, military engineering and mining. His industry experience includes engineering, maintenance and logistics management of high value fleets of equipment and assets and his consulting experience spans all phases of the asset management lifecycle from investment planning and strategy to operations and maintenance and disposal and divestment.

Oliver **Skelding,** B.A.(Economics), M.Ec (Econometrics)

Role: Econometrician

Oliver is a senior analyst in our Sydney office who has experience in the regulation of monopoly industries, economic benchmarking and the application of econometric techniques.

Relevant Skills

Knowledge of the regulatory framework within the National Electricity Market
Knowledge of Australian DNSP cost structures
Total factor and partial productivity analysis
Econometric modelling

Relevant Experience

Oliver has worked with a number of Australian DNSPs' to identify expenditure outcomes relative to other operators within the Australian electricity supply industry. Recent engagements include;

- Working with a Victorian DNSP to benchmark expenditure relative to other businesses in the NEM. This project involved using both the AER's benchmarking techniques and other available benchmarking techniques such as Data Envelopment Analysis.
- Assisting an Australian TNSP with benchmarking in preparation for its revenue proposal to the Australian Energy Regulator.
- Working with a number of DNSP's to highlight possible outcomes of the application of the AER's preferred benchmarking techniques.
- Developed performance reports and conducted performance analysis for a number of functions for a large infrastructure manager.
- Developed safety and risk analysis and reports for electrical safety incidents for a state safety regulator.

Professional Summary

Oliver has completed a Master of Economics, specialising in Econometrics. Prior to working with Huegin he worked for the NSW Department of Finance and Services.

At Huegin, Oliver has responded on behalf of Australian DNSPs to the Australian Energy Regulator's Better Regulation Paper regarding the difficulties of using econometric benchmarking techniques within the context of Australian DNSPs and TNSPs. Oliver has also assisted with the benchmarking of Australian DNSPs and TNSPs in preparation for revenue proposals to the Australian Energy Regulator.

Dr Ben **Petschel,** B.Sc.(Mathematics), PhD (Mathematical finance)

Role: Mathematician

Ben has extensive experience in mathematical modelling for cost optimisation and risk management for both multi-national organisations and academic institutions. Ben is highly experienced in benchmarking methodologies, techniques and analysis. He has worked in projects in the energy, transport, mining and financial services industries.

Relevant Skills

	Knowledge of regulatory requirements for electricity networks
	Financial services risk and governance
	Safety risk modelling expertise
	Fault tree analysis
	Complex mathematical modelling
	Statistical and mathematical analysis

Relevant Experience

Ben is an experienced mathematician who has worked in industry applying his skills to various cost, risk and asset management problems. In the past three years Ben has worked extensively in electricity and transport infrastructure management. His experience includes:

- Capital works program simulation for a large electricity network, where Ben modelled and analysed the workforce demand associated with various .
- Capital project portfolio optimisation, where Ben has built and deployed investment decision support models that allocate capital across projects based on advanced, complex decision algorithms for electricity and transport industry asset managers. .
- Post tax revenue model scenario modelling, where Ben has built and analysed probabilistic capital and operating expenditure profiles associated with various x-factor and consumer behaviour scenarios.
- Asset replacement forecasting for a large electricity distribution network.
- Developed and implemented a model for wooden electricity pole maintenance for the purposes of risk management and consideration of impacts to changes in inspection and maintenance programs of an electricity distributor

Professional Summary

Ben studied applied mathematics and mathematical finance at the University of Queensland. He has worked in Operational Risk modelling for St.George Bank and Westpac, and Risk Model Validation for ANZ Bank. His PhD thesis was on the weather impacts on derivatives trading prices. Since joining Huegin, Ben has designed and built many electricity industry investment and cost models, analysed asset forecasts and provided mathematical and analytical support to many distribution businesses.

Naomi **Donohue,** B.Bus (Accountancy and Computer Applications), CPA

Role: Regulatory Expert

Naomi is our Brisbane based senior manager. Naomi was involved in the AER's Better Regulation process and has expertise in distribution network service provider regulation and cost constructs. Naomi was key in unpacking the regulatory environment and cost breakdowns examined in the report.

Relevant Skills

	Regulatory Determination knowledge and experience
	Industry Operational knowledge
	Industry Regulation knowledge
	Benchmarking experience

Relevant Experience

Naomi is a qualified CPA with extensive experience in regulation and finance of electricity energy distributors. Naomi has specific expertise in the areas of regulatory determinations and national electricity market rule changes.

- Management and co-ordination of the financial related components of the revenue determination for an electricity distribution network service provider.
- Participation and involvement in the AER's Better Regulation program.
- In-depth understanding and knowledge of the energy regulation environment in Australia..
- Identification and strategic management of regulated and non-regulated revenue risks and opportunities, collation and presentation of expected costs for operations and infrastructure investment, and compliance with relevant national electricity law and regulatory requirements.
- Responsible for compilation and AER approval of a network service providers' Cost Allocation Model.
- Completion of all financial modelling to a support a network service providers' Regulatory proposal utilising AER models without any compliance or regulatory issues.

Professional Summary

Naomi has significant experience working in the regulated electricity sector, having worked in a distribution network service provider's regulatory and financial departments for over 8 years prior to joining Huegin. She is also experienced working with government agencies to achieve both commercial and social outcomes. Naomi is a qualified CPA with over 20 years experience in management accounting, strategic planning, process improvement and regulation.

Darryl **Walker,** B.Eng (Aero), MSc (Thermal Power)

Role: SME

Darryl is one of our SMEs on this project and is responsible for those areas that involve a detailed knowledge of either maintenance or risk. Darryl has been involved in most Huegin benchmarking studies over the past five years. Darryl also been working with electricity distributors in non-network areas including fleet, property and IT.

Relevant Skills

Industry benchmarking
Performance assessment and operational improvement
Stakeholder management, engagement and communication
Stakeholder communication
Maintenance
Risk modelling

Relevant Experience

Darryl is an experienced engineer and consultant with specific expertise in the areas of risk management and modelling as well as the management of large physical assets.

- Led teams in the modelling and performance improvement of capital works programs in electricity distribution organisations.
- Led numerous asset management assignments within electricity distribution organisations. For Ergon Energy this culminated in the development and implementation of the methodology and toolset for measuring Network Health.
- Recently led a team that developed a model to simulate the lifecycle of a wood power pole. this model, calibrated across Australia, provides an indication of the severity of the environment (from the perspective of pole degradation) and the likely failure probability. Further, this model allows asset managers to optimise their inspection and maintenance regimes.
- Led a team that worked with Ergon Energy to understand and analyse the risks associated with the network using the Dangerous Electrical Event data.

Professional Summary

Darryl Walker is a Director of Huegin Consulting. Darryl has 18 years postgraduate experience in both management and consulting with a focus on capital intensive industries including construction, development, airlines, transportation, banks and utilities.

Prior to joining Huegin in 2008, Darryl spent ten years in the Royal Australian Air Force as an engineering officer. After leaving the Air Force, Darryl spent time working for both Booz Allen Hamilton and KPMG in the areas of strategy and performance improvement. Recently Darryl has been working extensively with both government and commercial organisations to understand maintenance, asset management,risk and reporting on and improving performance.

Annex C - Bauer consistency tests: model specifications

C

Opex Partial Factor Productivity Models

The table below displays the different variables that have been used to produce output indexes for sensitivity analysis of the different efficiency scores obtained using different model specifications.

Outputs	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Energy Distributed (GWh)	✔	✔		✔	✔		✔	✔
Customer Numbers (Total)	✔	✔		✔	✔	✔	✔	✔
Residential Customers			✔					
Commercial Customers			✔					
Small Industrial Customers			✔					
Large Industrial Customers			✔					
Customer Minutes Off Supply	✔	✔	✔	✔		✔	✔	✔
Ratcheted Peak Demand (MW)		✔		✔	✔	✔	✔	✔
Circuit Kilometres				✔	✔			✔
Route Line Length							✔	
MVA-kms	✔							

Data Envelopment Analysis Models

Outputs	Model 1	Model 2	Model 3	Model 4
Energy Distributed (GWh)			✔	✔
Customer Numbers (Total)	✔	✔	✔	✔
Ratcheted Peak Demand (MW)	✔	✔	✔	✔
Circuit Kilometres	✔	✔	✔	✔
Returns to Scale Assumption	Constant Returns to Scale	Variable Returns to Scale	Constant Returns to Scale	Variable Returns to Scale

Econometric Modelling

The table below gives the estimates obtained when the effects of share of underground on technical efficiency score (Battese and Coelli 1995) are modelled[32] - this means that Share of Underground is not included as part of the cost function but is used to adjust the efficiency scores obtained.

	Estimate	Std.Error	Z-Value	Pr(>|z|)	
(Intercept)	-0.2083076	0.3207431	-0.6495	0.5160	
Ln Customer Numbers	0.6220192	0.0549831	11.3129	< 2.2e-16	***
Ln Circuit Length	0.1408193	0.0239694	5.8750	4.23E-09	***
Ln Ratcheted Max Demand	0.2230500	0.0532467	4.1890	2.80E-05	***
Year	0.0183342	0.0039190	4.6783	2.89E-06	***
Ontario	0.0430338	0.0393800	1.0928	0.2745	
New Zealand	-0.0021316	0.0457649	-0.0466	0.9629	
LNShareUGC	-0.1800206	0.0182320	-9.8739	< 2.2e-16	***
sigmaSq	0.0537836	0.0053514	10.0504	< 2.2e-16	***
gamma	0.6661210	0.0948595	7.0222	2.18E-12	***

[32] The frontier package (Coelli and Henningsen) has been used in R to produce the results

Contact Details

Jamie Blair

Level 10, 2 Elizabeth Plaza, North Sydney, NSW 2060

t: +612 9409 2400
e: jblair@huegin.com.au

huegin