Graeme Finlayson
General Manager AER
Ergon Energy PO Box 1090
Townsville QLD 4810

By email

23 July 2015

Dear Graeme

Peer review of Huegin report

1 My instructions

Ergon Energy (**Ergon**) has asked me to provide a peer review of, and opinion on, Huegin's report entitled *AER Benchmarking of Ergon Energy Opex: Huegin Review of the Preliminary Determination* dated 1 July 2015 (**Huegin report**). In addition to the report, I have been provided with several files containing computer code, computer output and spreadsheet calculations that underpin some of the results in the report.

In preparing this report, I have sought the advice of several colleagues at Frontier Economics (**Frontier**).

2 Summary of Huegin's key arguments

The key arguments made in the Huegin report can be summarised as follows:

- Ergon has submitted a proposed forecast of opex to the Australian Energy Regulator[1] (AER) that included considerable efficiency gains
- The AER has rejected Ergon's base year opex forecast and substituted its own forecast
- The AER's forecast is based on a statistical benchmarking model that has a number of limitations:
  - the validity of the model is based on assumptions that have not, and in some cases, cannot be validated
  - there is estimation uncertainty in the efficiency scores produced by the model that has not been adequately taken into account by the AER in determining its proposed base year opex
  - there are alternative models with equally good statistical properties that produce quite different outcomes
  - insufficient attention has been paid to Ergon's unique nature as a large, rural and remote network. Ergon is an outlier compared to all other networks in the sample used to estimate the AER's benchmarking model, making the model unfit for explaining Ergon's opex
  - insufficient weight has been given to the most recent years of data in which Ergon has achieved considerable opex savings.
- The AER's approach to making adjustments to the benchmarking results to account for operating environment factor (OEF) adjustments not captured by the model is blunt and ad hoc, with very little analysis to support the adjustments.
- In view of the limitations of the preferred statistical benchmarking model it is inappropriate for the AER to rely on the model to determine a single opex forecast. This is a form of model selection bias. To avoid this bias, a number of other regulators base their final efficiency assessments on the outputs of multiple models and techniques.
- Given the inherent uncertainties associated with the outcomes of the statistical modelling process, and the fact that Ergon's proposed opex forecast is only slightly above the upper confidence limit for the AER's preferred model, and is within the confidence limits for the AER's alternative least squares econometric (LSE) models, the AERs' rejection of Ergon's proposed opex forecast is not justified.

---

[1] The development of the AER's benchmark model and the base year opex forecasts was undertaken by Economic Insights (EI) on behalf of the AER. I will refer to EI's models results as the AER's models and results.

3 Comments

3.1 Overall assessment

Having reviewed the Huegin report and supporting material, and given my own experience in analysing the dataset used in the AER's modelling, I find the arguments presented in the Huegin report, and the evidence to support those arguments, persuasive. While I have some reservations about aspects of the report, these reservations do not detract from the main thrust of Huegin's critique of the AER's benchmarking exercise, and I concur with the main conclusions in the report. Below I elaborate on several aspects of Huegin's arguments, focusing in particular on data issues and the benchmark modelling.

3.2 Specific comments

3.2.1 Data from different jurisdictions are not poolable

The Huegin report (pp.13-16) provides overwhelming evidence that Ergon's network is quite unlike the networks in Ontario. The report argues that Ergon's service area is 70% larger than all of Ontario and that only one of the 37 Ontarian networks is not contained in an urban area. Any model of opex based on such an 'unmatched' sample is unlikely to provide a meaningly model of what drives costs in a large, rural, sparsely populated network.

This view is supported by formal statistical testing undertaken by Frontier, which overwhelmingly rejects the hypothesis that the data for Ontario can be pooled with Australian data.[2] I find Economic Insights' (EI) reasons for rejecting Frontier's statistical tests unconvincing.[3] Additional testing I have undertaken, for example, of the equality of the time trends across countries, that weights attached to the drivers of opex differ between the countries in the sample. Dummy variables the only shift the intercept of the cost function between countries are not sufficient to adequately capture those differences. As a result, the AER's preferred model will produce biased (and statistically inconsistent) predictions of opex.

---

2 Frontier Economics (Jan 2015), *Review of the AER's econometric benchmarking models and their application in the draft determinations for Networks NSW*, Section 3.2.3.

3 Economic Insights (22 April 2015), *Response to Consultants' Reports on Economic Benchmarking of Electricity DNSPs*, p.25.

3.2.2 Alternative models provide a better fit to the data

In Section 4 of the report, Huegin provides persuasive evidence why opex is likely to increase non-linearly with the network service area. This raises the question of whether the log-linear relationship assumed in the AER's preferred stochastic frontier analysis (SFA) Cobb Douglas (CD) model provides enough flexibility to capture the way drivers related network size and network density impact on opex.

Huegin have estimated a model with a squared term for (log) circuit length to provide more flexibility in the relationship between network size and opex. This model is statistically superior to the AER's preferred model on the Wald and likelihood ratio tests (Huegin report, p. 37). This confirms Frontier's own findings in regard to similar extensions of the SFA CD model.

Some extensions of the SFA CD model that have additional higher order terms (such as squares and cross-products of the variables) are statistically superior to the SFA CD model, and they can produce vastly different efficiency scores and forecasts of base year opex. EI dismissed one such model, the SFA translog (TL) model, on the grounds that this model results in some elasticities having the 'wrong' sign. I believe that EI's objections are unwarranted since these so-called violations are typically very minor and statistically highly insignificant.

I support Huegin's view that there are other statistical models that are, on statistical grounds, better than the AER's SFA CD model. Since some of these models produce quite different efficiency scores to the SFA CD model, it is, in my view, unrealistic for the AER to base its benchmarking efficiency scores solely on the SFA CD model. This totally ignores the uncertainty introduced into the benchmarking exercise through the model selection process.

I do not agree with Huegin that the preference of one model over another in this case is simply a matter of professional opinion (Huegin report, p. 27). That would be true if there were mixed results on different statistical criteria regarding the adequacy of the respective models. However, where one model provides a better fit to the data than another model on a range of widely accepted criteria, that model should be selected as the preferred model. In my view, the SFA CD model does not fit that description, whereas there are other models that do.

3.2.3 The AER pays insufficient attention to estimation uncertainty

The SFA CD model requires a large number of statistical assumptions to be satisfied for its validity. If these assumptions are not satisfied, the estimates of the model's parameters are likely to be statistically inconsistent, leading to biased estimates of the model's parameters and the resulting efficiency scores. The Huegin report mentions the requirement for the distribution of inefficiency scores to follow a truncated normal distribution (Huegin report

p.9). Among these assumptions are homoscedasticity, lack of autocorrelation and normality of the idiosyncratic error term. None of these assumptions has been tested and validated by the AER.

Even if all the assumptions required for a model such as the SFA CD are satisfied, the parameter estimates and estimated efficiency scores are subject to estimation error. Estimation error is usually indicated by the standard error of an estimate, or a confidence interval providing a range of values around the estimate.

The Huegin report (p.9) points out that, by relying on a single estimate of inefficiency, the AER pays insufficient regard to the uncertainty inherent in any econometric estimates. In Figure 5 of the report (p.31), Huegin presents confidence intervals around the base year opex predicted by the AER's preferred model SFA CD and its alternative LSE models. The figure shows that Ergon's proposed base year opex forecast lies within the confidence intervals for the LSE models, and quite close to the interval for the SFA CD model. Visually, this provides quite a different view of Ergon's efficiency compared to the raw score of 0.48 produced by the AER's SFA CD model.

While I have not been provided with the computer code or Excel files that would enable me to verify these results, I find this an effective means of conveying the uncertainty inherent in the estimation of efficiency scores, and a better starting point for a discussion on target opex reductions than a single efficiency score.

3.2.4 AER should have focused on most recent data

The Huegin report argues that the AER should have focused on more recent data rather than outdated historical data that fails to recognise recent improvements in efficiency made by Ergon (p.36). The report presents results for several versions of the AER's SFA CD model that only differ in the sample used to estimate the model (see, for example, models 8 to 13 in Figure 13 on p.35). These results show that the forecast of base year opex is quite sensitive to the estimation period.

While the report does not provide estimates of these efficiency gains, I note that in the dataset provided by the AER for the benchmarking exercise, Ergon's opex in 2012/13 was 16% less than in 2011/12 and 10% less than in 2010/11. This lends some weight to the assertion that Ergon has become more efficient in recent years. I concur with Huegin's view that the AER's methodology leads to estimated efficiency scores that could badly misrepresent a DNSP's current level of efficiency.

However, I do not agree with Huegin's suggestion that the AER should follow Ofgem's approach and make use of forecasts prepared by the DNSPs in its benchmarking analysis. Huegin includes such forecasts up to 2019 in its Model 12 and Model 13.

In this approach there is a risk that some DNSPs could seek to influence the outcome of the benchmarking exercise in their favour by submitting 'strategic' forecasts. One of the ways Ofgem mitigates the risk of such gaming is through a form of menu regulation, which provides incentives to businesses to forecast accurately, and to reveal those forecasts truthfully to the regulator. The AER's regulatory framework does not feature menu regulation, and hence it would be inadvisable to follow this approach in Australia.

3.2.5 Overseas regulators rely on a range of approaches

The above discussion indicates that any model aimed at determining efficiency scores for regulated businesses is subject to a range of assumptions and data issues, and the estimated efficiency scores are sensitive to how these issues are handled. Hence, in my view, it would be unwise to rely on a single approach and a single number to characterise the efficiency of a particular business.

The Huegin report cites the Productivity Commission as stating that benchmarking is not about identifying a single number denoting the efficiency of a business, but rather a potential range of likely numbers (p. 12). The Huegin report argues that regulators in other jurisdictions understand the dangers of relying on a single econometric model to determine efficiency scores for benchmarking purposes, and that, unlike the AER, they rely on a range of approaches when benchmarking businesses.

Pages 10 to 12 of the Huegin report discusses several examples of regulators who rely on range of approaches, including Ofgem and the German Bundesnetzagentur. There are some minor inaccuracies in Huegin's descriptions of the benchmarking approaches used by some of these regulators, but these minor inaccuracies do not undermine the key point made by Huegin that overseas regulators experienced in economic benchmarking use a number of different approaches to mitigate the risk that regulatory decisions are influenced excessively by model errors.

Please do not hesitate to contact me if you wish to discuss any of my findings in further detail.

Yours sincerely,

Professor Robert Bartels