huegin

Submission on the AER Expenditure Guidelines

A review of the benchmarking techniques proposed

Date: 20 September 2013

Author: Huegin Consulting Group

Version: 1.2

www.huegin.com.au

Contents

1

Introduction

A short overview of the new approach to benchmarking outlined in the AER Expenditure Guidelines of August 2013.

Introduction

On the 9th August 2013 the Australian Energy Regulator released its Expenditure Forecast Assessment Guideline (referred to hereafter as the Guideline) outlining the techniques it will use when evaluating the prudence and efficiency of Australian Network Service Provider (NSP) expenditure. The tools announced by the AER are a combination of existing techniques and new benchmarking techniques.

New Techniques

In an effort to reduce information asymmetry between the AER and individual NSPs and in response to changes in the National Electricity Rules (the Rules) the AER will now use more sophisticated benchmarking techniques when evaluating NSP expenditure, these include;

- Tornqvist Multilateral Total Factor Productivity (MTFP)
- Data Envelopment Analysis (DEA)
- Econometric analysis
- Category level benchmarking[1]

Existing Techniques

The AER will continue to utilise techniques used in previous NSP revenue determinations, these include;

- Economic justification for expenditure
- Expenditure forecasting methodology
- Reviewing expenditure governance and policies
- Trend analysis
- Category analysis
- Targeted review of projects and programs
- Sample review of projects and programs

An overview of how the techniques will be applied is provided on the following pages, whilst the remainder of this document contains commentary on the approach and the individual techniques introduced through the guidelines.

[1] High level category benchmarking has been used in previous determinations (ratios such as opex/km and capex/load density) however the guidelines suggest that category benchmarking will now be conducted in much greater detail than previously.

2

The Assessment Approach

An illustration of the approach structure and commentary on the approach from our review of the Expenditure Forecast Guideline.

The Structure of the Approach

The table below illustrates how we believe the new benchmarking techniques will be used by the AER.

Expenditure Category	Benchmarking technique	Outcome
Total expenditure	MTFP	• Overall efficiency and rate of change in efficiency • Growth of inputs and outputs • Forecast future TOTEX
	DEA	• Cross check results of MTFP
Capital expenditure	Category benchmarking	• Augmentation CAPEX • Replacement CAPEX • Non-network CAPEX • Customer initiated CAPEX
Operating expenditure	MTFP	• High level indication of OPEX efficiency
	Econometric analysis	• Base year efficiency • Annual rate of change
	Category benchmarking	• Maintenance and emergency response OPEX • Vegetation management OPEX • Overheads

The following sections provide commentary on the overall approach with specific commentary on the individual techniques in the next chapter.

Observations

The following sections highlight some observations and concerns regarding the overall approach detailed in the Guideline. Our most significant concern is that there remains a high degree of uncertainty in the expected outcomes of the approach, reflected by the lack of specification of the models to be used. It appears that the model specification will be guided by the available data - if this is the case, it is unlikely that the model results will be reliable, particularly in the initial years of application.

The approach - holistic or selective?

The AER has suggested that their approach to benchmarking is holistic, the inference being that the introduction of several techniques somehow provides a more robust evaluation framework. In reality, the use of several techniques is more likely to lead to selection bias in the results. When presented with several techniques that rely on different principles and produce different results, the tendency is for the practitioner to favour the techniques that best fits the conclusions they are trying to draw. This tendency is compounded when the modelling relies on using the data to test the "fit" of the models, rather than proper hypothesis testing. That is, when models are selected through searching for patterns in the available data, rather than through the declaration and subsequent testing of objective hypotheses, the user can very quickly lose sight of the practicality of the model application.

The approach - complementary or competing?

The Guideline infers that several of the techniques introduced are complementary; this follows previous guidance from the AER earlier in 2013 that where modelling techniques have weaknesses and strengths in different areas, they can be applied in a complementary fashion. This is not the case. The errors in one model do not cancel out the errors in another. The different techniques selected for inclusion by the AER have fundamentally different technical origins and characteristics; they rely on different assumptions. For example, DEA is often used for convenience, rather than preference; the advantage of DEA is that the cost production function does not need to be known. The DEA model finds a best fit amongst a data set of input and output combinations, thereby assuming a cost production function based on the sample data. Econometric modelling does the opposite - it relies on the upfront declaration of a cost production function for the industry and then calculates a residual for each business against the industry average. Combining DEA and econometric modelling suggests that the cost production function is both unknown and known.

Undue reliance on the theory

Efficiency frontiers and industry cost production functions are economic concepts only; neither actually exist. Whilst useful for modelling purposes, there are two significant issues with relying on these models to support actual cost adjustments:

1. The validity of the models and veracity of the results they produce is poor due to the statistically insignificant sample size; and

2. The random, experimental nature in which the models are constructed - by finding the model that best fits the available data - leads to a model of best fit *at that point in time,* rather than a robust, one-size-fits-all formula for any industry participant.

All of the techniques introduced by the AER have been used in electricity distribution regulation across a broad range of jurisdictions. In most cases, the techniques have been introduced, abandoned and sometimes reintroduced over time as the models either fail to stand up to repeat application, fail to produce the intended industry consequences or simply fall in and out of favour. Many of the techniques require many years worth of data, but few last long enough to achieve it. At the very least, the model specifications change over time as new data arrives with each regulatory period. Whilst this might seem like simple refinement of models, it has significant consequences for the businesses who are forced to react to the signals produced by the models - often in tension with the signals they are still adjusting to from the previous incarnations.

Even in Europe, where many of these techniques have been more thoroughly tested in a much more suitable environment (larger sample sizes and more homogenous conditions), practical application remains challenging.

> ***"in practice, benchmarking has proven either troublesome or irrelevant to the regulatory process, but proponents continue to search for "better" models that will be more useful."***[2]

> ***"Our view is that the top down approach combines costs to a degree that relatively simple cost drivers are unable to identify and differentiate between the differing circumstances of the DNOs."***.

The UK regulator, OFGEM, is an example of a regulator who has tested the veracity of their benchmarking approaches over time and subsequently changed the approach numerous times in the last decade. A case study on OFGEM's latest thinking is included in Chapter 4.

Application of benchmarking

Huegin advocates and applies benchmarking as an informative process for identifying and communicating differences in the cost outcomes between businesses. We see the results of benchmarking as the means for initiating investigations into productivity and efficiency improvement opportunities, not the ends. In our experience, those investigations invariably uncover another level of detail about cost drivers. It takes considerable time and effort to determine the root cause of benchmarked cost differences and similar effort again to determine the ability of the business to influence those causes. Our experience is that benchmarking is not reliable in predicting an industry cost function and should not be used as a substitute for forecasts.

[2] Shuttleworth, G. NERA pg 310 Benchmarking of electricity networks: Practical problems with its use for regulation

3

The Techniques

An overview of the techniques and an evaluation of their attributes, applicability and the assumptions that underpin them individually and collectively.

The Techniques

The following sections provide an overview of the characteristics and some issues with the new techniques introduced through the Guideline.

Tornqvist Multilateral Total Factor Productivity (MTFP)

The AER have indicated that they will use MTFP to benchmark total expenditure and use the results to measure the change in and overall efficiency of NSPs, the historic growth of inputs and outputs and to forecast future aggregate expenditure.

Two problems with using MTFP to infer efficiency are that MTFP is unable to account for the influence of different business conditions on efficiency results and it is also unable to account for the influence of scale on efficiency results. These were highlighted in a 2003 report conducted by CEPA for OFGEM.

> ***"It provides only limited ability to control for differences in the business environments of firms in the sample group."***
>
> ***"the approach is unable to distinguish scale effects from efficiency differences."***

Given operating environments within the National Electricity Market (NEM) are far more varied than those within the United Kingdom, it is unlikely that the model outlined by the AER will be able to overcome these shortcomings of using MTFP to infer efficiency between networks.

The outputs in the AER's preferred model are number of customers, system capacity and reliability. If differences in MTFP results are used to infer efficiency between NSPs then an assumption that needs to be made is that these different MTFP scores are due solely to business conditions that NSPs are able to control. Given customer connections and system capacity are two outputs used in the preferred MTFP model one way a NSP can appear more efficient is by increasing the size of its network or number of customers relative to its index of inputs. For a rural network, such as Ergon Energy, the cost drivers associated with increasing network size and its customer base include low customer density (on average Ergon has to build longer line lengths per customer) and customer demographics (Ergon has the second highest average consumption levels per customer). By contrast, if an urban network, such as Ausgrid's, outputs in terms of customer connections and system capacity were to increase its cost drivers are more likely to be network design (much of its network is underground and high voltage), high reliability standards and asset age. These differences are not confined to Ergon and Ausgrid, there are different combinations of cost drivers between all NSPs in the NEM that make a simple MTFP comparison using three outputs and four inputs misleading and unreliable.

In response to the difficulties of using Tornqvist MTFP to benchmark NSP expenditure the AER have indicated that a second stage regression will be used on MTFP results to remove scale effects and exogenous business characteristics. Given a sample size of 13 networks it is unlikely that the dataset will be large enough to remove differences in MTFP results due to uncontrollable business conditions, scale differences and data errors and to infer relative level of efficiency between NSPs.

Data Envelopment Analysis

The AER have indicated they will apply DEA as a cross-check of the MTFP results. This infers that if the relative ranking of an NSP through both MTFP and DEA is similar, then the AER will assume that the analysis is sound. MTFP and DEA are two separate techniques, but both are sensitive to the model specification. It is possible to produce entirely contrary results from MTFP and DEA with the same dataset. It is also possible to produce similar results through coincidence or through poor model specification. Our concern is that in the case of contrary rankings the results will be discarded, whereas if the models provide similar rankings then the results will be endorsed and used.

DEA as a standalone technique is susceptible to small sample sizes and heterogenous data. Australian electricity networks carry both of these characteristics. DEA is more suited to other industries where business conditions are transferrable (e.g. a factory), the markets are competitive and businesses have full control over inputs and there are many businesses (called Decision Making Units, DMUs, in DEA) to compare. The European electricity networks, where there are many smaller networks over a smaller, more homogenous environment, are somewhat more suited to the application of DEA. However even in the European environment, issues have been encountered with the heterogeneity of conditions. In DEA such heterogeneity is often addressed by clustering DMUs into similar groupings - thereby producing multiple frontiers by attribute group. This is not possible in Australia as it would reduce an already insignificant sample size into groups of only two or three businesses.

Econometric analysis

The AER have indicated they will use an industry opex cost model as a counterfactual to compare the efficiency of a NSPs base year opex and also future changes in opex.

Using Ordinary Least Squares (OLS) to estimate an industry short run cost function means finding a line of best fit between different observed NSP operational expenditures. The problem with applying this technique is that there must be sufficient explanatory variables in the model to account for the different characteristics that drive the differences between NSP operational expenditures. This is unlikely given a limited dataset and significant differences in NSP operating environments.

The variable used to compare businesses in an econometric model is the residual. The danger with using the residual from an econometric model (the difference between modeled and actual cost) to interpret base year efficiency is that the components of this residual are likely to include statistical noise, data errors and unexplained business conditions. The residual should be interpreted as variation between the predicted opex cost and actual cost that remains unexplained by the model and not just as inefficiency.

Having so many different drivers of opex costs between networks there is also a chance that the opex model constructed by the AER will be influenced by omitted variable bias. The following example is a simple opex cost function in which opex costs are driven by customer connection, reliability and system capacity, capital inputs (line length) and an exogenous variable (network design which could be the percentage of a network underground) and time (technology change):

$Opex_{it} = B_1$Customer Connections + B_2Reliability + B_3System Capacity + B_4Line Length + B_5Network design + B_6time + u_{it}

Given the many different influences on opex cost it is easy to think of different omitted explanatory variables that could lead to biased results. One example is asset age, by not including it within this opex model it ends up in the residual. If there is a relationship between asset age and reliability, which seems likely, then the coefficient B_2 will be biased.

The AER have proposed using the opex cost function derived to determine the efficiency of base year opex and consequently to determine future changes to opex productivity. This technique involves using output and input elasticities, time trends and business condition variables from the opex cost function. Using these elasticities is likely to overstate the economies of scale that can be achieved by different businesses.

$$P\dot{F}P_{OM} = (1-\sum_{i} \varepsilon_{i\,Y_i}) \cdot \dot{Y}^{\varepsilon} - \sum_{m} \varepsilon_{X_{K,m}} \cdot \dot{X}_{K,m} - \sum_{n} \varepsilon_{Z_n} \cdot \dot{Z}_n - \dot{g} - \dot{\eta}$$

The formula above is the technique proposed by the AER in determining the partial productivity of operating expenditure beyond the base year. As discussed above there are a number of reasons to believe that the coefficients in this model will not be representative of the true values for individual networks. For a rural network, the opex cost elasticity associated with increasing output is likely to be higher than the estimated industry opex cost elasticity.

Category analysis

The AER have indicated they will conduct category benchmarking for different categories of opex and capex when determining the efficiency and prudence of NSP expenditure.

We believe that category analysis cannot be used to indicate a NSPs level of efficiency for the same reasons that OLS regressions do not work. There is not one common cost driver for the different categories proposed by the AER. It is difficult to obtain any meaningful interpretation of efficiency in category benchmarking because costs are rarely driven by a sole driver and those drivers never exhibit the same influence across different businesses. This example highlights the difficulties in benchmarking heterogenous networks; there is no common cost driver among networks that allows a consistent comparison of expenditure. This may also lead to the AER choosing metrics that make a particular NSP appear inefficient whilst ignoring other metrics that make them appear efficient.

Challenging the approach assumptions

The individual techniques and the overall approach outlined in the Guideline rely on many assumptions - some stated, some implicit - which should be tested for defensibility. We believe through our own testing and experience, that several of the assumptions are not justified and that the error and uncertainty inherent in the approach is more significant than the fidelity of the adjustments that will result from the application.

Issues that apply to one or more of the techniques proposed by the AER include:

1. Model specification - the inputs and outputs of an electricity network have been a topic of debate for a considerable amount of time; given the sensitivity of the models to the specification, results must be considered in the context of appropriate error tolerances.

2. Data heterogeneity - business structures, ownership differences, accounting differences and variations in the scope of responsibilities (for example, division of responsibility between NSPs and councils for vegetation management and public lighting) are all significant impediments to homogenous data sets.

3. Sample size - 13 NSPs is quite simply insufficient to support economic modelling of appropriate statistical significance.

4. Exogenous factors - Australian NSPs face a complex set of cost drivers that affect NSPs in different ways. In our experience, the variation in costs due to variation in operating conditions is just as diverse within a single NSP (particularly those covering a majority of the state) as they are across industry. It is difficult to infer systematic inefficiency in this case, where a single business is clearly achieving a spectrum of cost outcomes that stretch well beyond its overall average.

In summary, our main concerns are with the two untested, but seemingly accepted, hypotheses inherent in the AER approach, which are:

1. Australian NSPs are comparable. That is, it is possible to compare Australian networks on a like-for-like basis and there is an industry cost function that represents all NSPs in the NEM.

2. The difference between the estimated costs and actual NSP costs are due to inefficiency and not other unexplained variables, different cost structures, data error or random error.

In our experience, neither of these hypotheses hold true.

The case study in the next chapter highlights the approach of OFGEM to the treatment of heterogenous businesses - in an environment much less variable than that in Australia.

4

A Case Study

The benchmarking challenge in the UK and the recognition by OFGEM of the influence of regional factors, challenges of economic benchmarking techniques and limitations of historical data.

Benchmarking opex in the United Kingdom

Benchmarking operational costs was a tool used by OFGEM for Distribution Price Control Review 4 (DPCR 4) and DPCR 5, which covers the period between 2005/06-2014/15. OFGEM has since announced a new approach to price determinations that gives limited attention to historic data.

> ***"There should be less emphasis on ex-post benchmarking of historic costs and a greater emphasis on benchmarking future plans, assessing the extent to which they represent value for money for customers."***[3]

OFGEM's approach to benchmarking in the next regulatory determination will change from being used to set base price operational expenditures to provide an overview of total cost efficiency in an "initial sweep" of NSP expenditure. This new approach was highlighted in the OFGEM's *Handbook for implementing the RIIO model.*

> ***"Total costs should be the basis of assessment given the ambition to avoid biasing the network company into particular solutions (e.g. capex over opex)."***[4]

> ***"We do not expect to use total cost benchmarking in a mechanistic analysis of the base revenue requirement given potential concerns about the robustness of the analysis".***[5]

Whilst using the same technique to benchmarking base year opex, OFGEM in DPCR4 and DPCR5 used a different process to that proposed by the AER. This process involved normalising opex cost data to a point where like-for-like comparisons could be made between NSPs.

Notably, there was a five stage process in determining base year opex efficiency that involved adjusting NSP cost data. These steps are illustrated below[6];

1. Reviewing the cost and efficiencies achieved by Distribution Network Operators (DNOs) during the existing price control period and their projected efficiencies for the next period;
2. Developing "normalised" and comparable cost information using actual costs;
3. Comparing actual normalised costs, using topdown benchmarking to help estimate efficiency cost levels;
4. Considering other information on efficiency; and
5. Adding back other cost items estimated separately to give the final opex allowance.

In normalising the opex costs before using benchmarking analysis, of particular consideration was the influence of regional factors. These were network sparsity for Scottish Hydro (SSE Hydro) and urbanity for London Power Networks (LPN). The following table summarises the results from DPCR 4 and DPCR 5 in relation to regional factors.

[3] pg2, RPI-X@20: The future role of benchmarking in regulatory reviews May 2010

[4] pg 64, Handbook for Implementing RIIO Model, OFGEM

[5] pg 64, Handbook for Implementing RIIO Model, OFGEM

[6] pg 65, Electricity Distribution Price Control Review Final Proposals, 2004

Control Period	Applicable Reference	Outcome
DPCR 4	"OFGEM considered that operating cost conditions were broadly similar for all companies with the exception of LPN and SSE Hydro. It is commonly recognised that employment costs are higher in London. SHEPD (SSE-Hydro) has a very large sparsely populated territory and as a result, incurs additional operating costs" pg 68. Electricity Distribution Price Control Review: Final Proposals, OFGEM, 2004	An allowance of £1.6 million per year for SSEs operating costs.
DPCR 5	"OFGEM considered that additional operating costs were incurred to service the SSE Hydro area and therefore made an adjustment to normalise costs" pg65-68, Electricity Distribution Price Control Review Final Proposals -Allowed Revenue - Cost Assessment, 2009	An increased allowance of £2 million per year.

When conducting opex benchmarking for the past two price determinations, OFGEM has recognised that there are additional opex costs associated with operating sparsely populated networks and highly urbanised networks, and has normalised the opex cost data before conducting econometric analysis to ensure like-for-like comparisons can be made. In the context of benchmarking Australian NSPs using the same technique (econometric analysis), normalisation of the input data would be far more challenging given the broader spectrum of conditions. The table below highlights the heterogeneity of the Australian networks compared to those in the United Kingdom.

Ratio	UK DNOs	NEM
*Difference in scale between the **longest network to the shortest***	3 times the size	39 times the size
*Difference in scale between the **largest and smallest, measured by customer numbers***	5 times the size	22 times the size
*Difference in customer density between the **most dense and least dense network***	4 times the density	12 times the density

The significance of these differences is amplified by the fact that OFGEM found it necessary to normalise the data even with a much lower heterogeneity threshold, whereas the AER approach assumes that the exogenous factors can be controlled for within the model itself.